

FIGURE 1

CCAATCGCCCGGTGCGGTGGTGCAGGGTCTCGGGCTAGTCA**ATG**GCGTCCCCGTCTCGGAGAC
TGCAGACTAAACCAGTCATTACTTGTTTTCAAGAGCGTTCTGCTAATCTACACTTTTATTTTC
TGGATCACTGGCGTTATCCTTCTTGCAAGTTGGCATTGTTGGGGCAAGGTGAGCCTGGAGAATTA
CTTTTCTCTTTTAAATGAGAAGGCCACCAATGTCCCCTTCGTGCTCATTGCTACTGGTACCG
TCATTATTCTTTTGGGCACCTTTGGTTGTTTTGCTACCTGCCGAGCTTCTGCATGGATGCTA
AAACTGTATGCAATGTTTTCTGACTCTCGTTTTTTTTGGTCTGAACTGGTCGCTGCCATCGTAGG
ATTTGTTTTTCAGACATGAGATTAAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGC
AGTATAACTCTACAGGAGATTATAGAAGCCATGCAGTAGACAAGATCCAAAATACGTTGCAT
TGTTGTGGTGTCAACGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGATT
TCCTAAGAGTTGCTGTAAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACA
ATGAAGGTTGTTTTATAAAGGTGATGACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGA
ATTTCCCTTTGGAGTTGCTTGCTTCCAACCTGATTGGAATCTTTCTCGCCTACTGCCWCTCTCG
TGCCATAACAAATAACCAGTATGAGATAGTG**TAA**CCCAATGTATCTGTGGGCCTATTCCTCT
CTACCTTTAAGGACATTTAGGGTCCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAACTG
ACAACACTACTTACTGATAGACCAAAAACTACACCAGTAGGTTGATTCAATCAAGATGTAT
GTAGACCTAAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGCAGTGGGCTGAT
TCAATCAAGATGTATGTTTGCTATGTTCTAAGTCCACCTTCTATCCCATTTCATGTTAGATCG
TTGAAACCCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAAATTGTAAATGAAGT

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA19902

><subunit 1 of 1, 245 aa, 1 stop, 1 unknown

><MW: -1, pI: 8.36, NX(S/T): 1

MASPSRRLQTKPVITCFKSVLLIYTFIFWITGVILLAVGIWGKVSLENYFSLLNEKATNVPF
 VLIATGTVIILLGTFGCFATCRASAWMLKLYAMFLTTLVFLVELVAAIVGFVFRHEIKNSFKN
 NYEKALKQYNSTGDYRSHAVDKIQNTLHCCGVTDYRDWTDTNYYSEKGFPSCKLEDCTPQ
 RDADKVNNEGCFIKVMTIIESEMGVVAGISFGVACFQLIGIFLAYCXSRITNNQYEIV

Important features of the protein:

Signal peptide:

amino acids 1-42

Transmembrane domains:

amino acids 19-42, 61-83, 92-114, 209-230,

N-glycosylation site.

amino acids 134-138

Tyrosine kinase phosphorylation site.

amino acids 160-168, 160-169

N-myristoylation site.

amino acids 75-81, 78-84, 210-216, 214-220, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 69-80, 211-222

FIGURE 3

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACAAA
GGAGTCGCGCCGCGCCGCGCCGCCCCCTCCCTCCGGTGGGCCCGGGAGGTAGAGAAAGTCAGT
GCCACAGCCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT
TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG
GCCGGGGTAGGCTCTGGAAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA
GCCGAGAGGTTTTCCACCGAGGCCCGCGCTTGAGGGATCTGAAGAGGTTCCTAGAAGAGGGT
GTTCCCTCTTTCGGGGGTCTCACCAGAAGAGGTTCTTGGGGGTCGCCCTTCTGAGGAGGCT
GCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG
GGAATAAGCTCTGCAACTTTCTTTGGCATTTCAGTTGTTAAAAACAAATAGGATGCAAATTCC
TCAACTCCAGGTTATGAAAACAGTACTTGAAAACTGAAAACTACCTAAATGATCGTCTTTG
GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGGTCTGTTGTTGACTCTCGAAGAG
CACATAGCCCACTTCCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCTGTATCTG
CCGAGATGACAGTGGAACAGATGACAGTGTTGACACCCAACAGCAACAGGCCGAGAACAGTG
CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTTCGGCCACCAAGGAGGGGC
CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTTGGACACACT
GGCAGTAATACGGACTCTTGTAAGATAAGTAAAGTATCTGACTCACGGTCACCTCCAGTGGAAT
GAAAAGTGTTCTGCCCCGAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG
CCAAGCCTTGCTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAAA
TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC
TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

FIGURE 4

MIVFGWAVFLASRSLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ
AENSAVPTADTRSQPRDPVRP PRRGRGPHEPRRKKQNV DGLVLDTLAVIRTLVDKO

Signal peptide:

amino acids 1-16

Casein kinase II phosphorylation site.

amino acids 22-26, 50-54, 113-117

N-myristoylation site.

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 5

GGCACGAGGCGCTGTCCACCCGGGGGCGTGGGAGTGAGGTACCAGATTGAGCCCATTTGGCC
CCGACGCCTCTGTTCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCCCTAACGGACTG
CAAG**ATG**GAGGAAGGCGGGAACCTAGGAGGCCTGATTAAGATGGTCCATCTACTGGTCTTGT
CAGGTGCCCTGGGGCATGCAAATGTGGGTGACCTTCGTCTCAGGCTTCCTGCTTTTCCGAAGC
CTTCCCCGACATACCTTCGGACTAGTGCAGAGCAAACCTCTTCCCCTTCTACTTCCACATCTC
CATGGGCTGTGCCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCA
CATTCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCCCTGAGCCTTACGCTGGCCACTGTCAAC
GCCCCTGGCTGGAACCCCGCACACAGCTGCCATGTGGGCCCTGCAAACCGTGGAGAAGGA
GCGAGGCCTGGGTGGGGAGGTACCAGGCAGCCACCAGGGTCCCGATCCCTACCGCCAGCTGC
GAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCTTCCGCTACCATGGGCTGTCC
TCTCTTTGCAATCTGGGCTGCGTCCTGAGCAATGGGCTCTGTCTCGCTGGCCTTGCCCTGGA
AATAAGGAGCCTC**TAG**CATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAA
AAAAAAAAAAAA

FIGURE 6

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107

<subunit 1 of 1, 231 aa, 1 stop

<NX(S/T): 0

MEEGGNLGGLIKMVHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFPFYFHISM
GCAFINLCILASQHAWAQLTFWEASQLYLLFLSLTLATVNARWLEPRTTAAMWALQTVEKER
GLGGEVPGSHQGPDPYRQLREKDPKYSALRQNFFRYHGLSSLCNLGCVLSNGLCLAGLALEIRSL

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 86-103, 60-75

Casein kinase II phosphorylation site.

amino acids 82-86

Tyrosine kinase phosphorylation site.

amino acids 144-151

N-myristoylation site.

amino acids 4-10, 5-11, 47-53, 170-176, 176-182

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 54-65

G-protein coupled receptors proteins.

amino acids 44-85

FIGURE 7

AATTCAGATTTTAAGCCCATCTGTCAGTGGAATTTTCATGAACTAGCAAGAGGACACCATCTT
 CTTGTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGGAAAAATGCTCTTTTGGGT
 GCTAGGCCTCCTAATCCTCTGTGGTCTTCTGTGGACTCGTAAAGGAAAACCTAAAGATTGAAG
 ACATCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCC
 AGAACTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAAC
 AGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAG
 AGAATGTCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGT
 CTGATCAATAATGCTGGTGTTCCTCGGCGTGTGGCTCCCACTGACTGGCTGACACTAGAGGA
 CTACAGAGAACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTC
 CTTTGGTCAAGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCA
 ATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAAGGTTTCAATGACAGCTTAAG
 ACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAA
 ACTTGGCAGATCCAGTAAAGGTAATTGAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCA
 GACATCAAACAACAATATGGAGAAGGTTACATTGAAAAAGTCTAGACAACTGAAAGGCAA
 TAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAA
 GTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAATTTTCTGGATACCTCTG
 TCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAA
 TCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGAT
 TTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGA
 TCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTG
 CTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGT
 ATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATG
 ATCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTT
 AAGTATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAA

FIGURE 8

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406

><subunit 1 of 1, 319 aa, 1 stop

><MW: 35227, pI: 8.97, NX(S/T): 3

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLT
 ESGSTALKAETSERLRRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDW
 LTLEDYREPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGF
 NDSLRRDMKAFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEgyIEKSLD
 KLKGNKSYVNMDLSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQK
 AELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 9

GCGGGCTGTTGACGGCGCTGCG**ATG**GCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTT
 CCTCTCAGTCGGACTTCCTGACGCCGCCAGTGGGCGGGCCCCCTTGGGCGGTGCGCCACCACT
 GTAGTCATGTACCCACCGCCGCCGCCGCCCTCATCGGGACTTCATCTCGGTGACGCTGAG
 CTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGGCGGCTCGTGCTGGAGGAAATGGA
 AGCAACTGTCGAGATTGCAGCGGAATATGATTCTCTTCCTCCTTGCTTTCTGCTTTTCTGT
 GGACTCCTCTTCTACATCAACTTGGCTGACCATTGGAAAGCTCTGGCTTTCAGGCTAGAGGA
 AGAGCAGAAGATGAGGCCAGAAATTGCTGGGTAAACAGCAAATCCACCCGTCTTACCAG
 CTCCTCAGAAGGCGGACACCGACCCTGAGAACTTACCTGAGATTTCTGTACAGAAAGACACAA
 AGACACATCCAGCGGGGACCACCTCACCTGCAGATTAGACCCCAAGCCAAGACCTGAAGGA
 TGGGACCCAGGAGGAGGCCACAAAAGGCAAGAAGCCCCCTGTGGATCCCCGCCCCGAAGGAG
 ATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGGCACCGAG
 CTCCCTTCAAGAAGAGCAGAAGTGCCACCAAGCCTCCCTGCCACCGGCCAGGACACAGGG
 CACACCAGTGCATCTGAACATATCGCCAGAAGGGCGTGATTGACGTCTTCCTGCATGCATGGA
 AAGGATACCGCAAGTTTGCATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCCTTCAGT
 GAGTGGTTTGGCCTCGGTCTCACACTGATCGACGCGCTGGACACCATGTGGATCTTGGGTCT
 GAGGAAAGAATTTGAGGAAGCCAGGAAGTGGGTGTGGAAGAAGTTACACTTTGAAAAGGACG
 TGGACGTCAACCTGTTTGAAGACAGATCCGCATCCTGGGGGGGCTCCTGAGTGCCTACCAC
 CTGTCTGGGGACAGCCTCTTCCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGC
 CTTCAGAACACCATCCAAGATTCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCACC
 CGCCACGGTGGACCTCCGACAGCACTGTGGCCGAGGTGACCAGCATTACAGCTGGAGTTCCGG
 GAGCTCTCCCGTCTCACAGGGGATAAGAAGTTTCAGGAGGCAGTGAGAAAGTGACACAGCA
 CATCCACGGCCTGTCTGGGAAGAAGGATGGGCTGGTGCCCATGTTTCATCAATACCCACAGTG
 GCCTCTTACCCACCTGGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTAC
 CTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGA
 AGCCATCGAGGGTGTGAGAACGCACCTGCTGCGGCACTCCGAGCCAGTAAGCTCACCTTTG
 TGGGGGAGCTTGCCACGGCCGCTTCAGTGCCAAGATGGACCACCTGGTGTGCTTCCTGCCA
 GGGACGCTGGCTCTGGGCGTCTACCACGGCCTGCCCGCCAGCCACATGGAGCTGGCCCAGGA
 GCTCATGGAGACTTGTTTACCAGATGAACCGGCAGATGGAGACGGGGCTGAGTCCCAGATCG
 TGCACTTCAACCTTTACCCCCAGCCGGGCGCTCGGGACGTGGAGGTCAAGCCAGCAGACAGG
 CACAACCTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCGTCACAGGGGA
 CCGCAAATACCAGGACTGGGGCTGGGAGATTCTGCAGAGCTTCAGCCGATTCACACGGGTCC
 CCTCGGGTGGCTATTCTTCCATCAACAATGTCCAGGATCCTCAGAAAGCCCGAGCCTAGGGAC
 AAGATGGAGAGCTTCTTCTGGGGGAGACGCTCAAGTATCTGTTCTTGCTCTTCTCCGATGA
 CCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCCACCCTCTGCCTATCT
 GGACCCCTGCC**TAG**GGTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAGAGGCACCTTG
 CTGGGTCTGTGGCATTTCCTCAAGGGCCACGTAGCACCGGCAACCGCCAAGTGGCCCAGGCT
 CTGAACTGGCTCTGGGCTCCTCCTCGTCTGCTTTAATCAGGACACCGTGAGGACAAGTGA
 GGCCGTGAGTCTTGGTGTGATGCGGGGTGGGCTGGGCCGCTGGAGCCTCCGCCTGCTTCCTC
 CAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAGGTCTCTGTGGGCCGACCA
 GAGGGGGGCTTCGAGGTGGTCCCTGGTACTGGGGTGACCGAGTGGACAGCCCAGGGTGCAGC
 TCTGCCCGGGCTCGTGAAGCCTCAGATGTCCCAATCCAAGGGTCTGGAGGGGCTGCCGTGA
 CTCCAGAGGCCTGAGGCTCCAGGGCTGGCTCTGGTGTGTTACAAGCTGGACTCAGGGATCCTC
 CTGGCCGCCCCGCAGGGGGCTTGGAGGGCTGGACGGCAAGTCCGTCTAGCTCACGGGCCCCCT
 CCAGTGAATGGGTCTTTTCGGTGGAGATAAAAGTTGATTGTCTTAACCGCAA

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529

><subunit 1 of 1, 699 aa, 1 stop

><MW: 79553, pI: 7.83, NX(S/T): 0

MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPPHRDFISVTLSFGESYDN
SKSWRRRSCWRKWKQLSRLQRMILFLLAFLLFCGLLFYINLADHWKALAFRLEEEQKMRPE
IAGLKPANPPVLPAPQKADTDPENLPEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTQEEAT
KRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSRRAEVPTKPPLPPARTQGT PVHLNY
RQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTMWILGLRKEFEEA
RKWVSKKLHFEKDVDVNLFESTITIRILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKI
PYSDVNIGTGVAHPPRWTSDSTVAEVT SIQLEFREL SRLTGDKKFQEAVEKVTQHIHGLSGK
KDGLVPMFINTHSGLFTHLGVFTL GARADSYEYLLKQWIQGGKQETQLLEDYVEAIEGVRT
HLLRHSEPSKLT FVGELAHGRFSAKMDHLVCFLPGTLALGVYHGLPASHMELAQELMETCYQ
MNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLRLPETVESLFYLYRVTGDRKYQDWG
WEILQSFSRFRTRVPSGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDDPNLLSLD
AYVFNTEAHPLPIWTPA

Important features of the protein:

Transmembrane domain:

Amino acids 21-40 and 84-105 (type II)

FIGURE 11

GGCGCCGCGTAGGCCCGGGAGGCCGGGCCGGGCTGCGAGCGCCTGCCCCATGCGCCGC
CGCCTCTCCGCACG**ATG**TTCCCCCTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGC
AGGTCCGGGTGCTCTCCGGCGGCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTTCGTGGC
CTGCCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGG
CCCGGGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCAGAG
CCGCCCCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCACCGCCTGGCAGTGCTGGT
GCCCTTCCGCGAACGCTTCGAGGAGCTCCTGGTCTTCGTGCCCCACATGCGCCGCTTCCTGA
GCAGGAAGAAGATCCGGCACCACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAAC
CGGGCAGCGCTCATCAACGTGGGCTTCTTGAGAGCAGCAACAGCACGGACTACATTGCCAT
GCACGACGTTGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCCTGAGGCTGGGC
CCTTCCACGTGGCCTCCCCGGAGCTCCACCCTCTCTACCACTACAAGACCTATGTGCGCGGC
ATCCTGCTGCTCTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGG
CTGGGGCCGCGAGGACGACGAGTTCTACCGGCGCATTAAGGGAGCTGGGCTCCAGCTTTTCC
GCCCCCTCGGGAATCACAACCTGGGTACAAGACATTTGCCACCTGCATGACCCAGCCTGGCGG
AAGAGGGACCAGAAGCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGG
AGGCCTGAACACTGTGAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCC
CCTGCACTGTCTCAACATCATGTTGGACTGTGACAAGACCGCCACACCCTGGTGCACATTC
AGCT**TGA**GCTGGATGGACAGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGA
CAAGGCCTCAGGTCGTGGGCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCA
AGCTACGCAATTGCAGCCACCCGGCCGCCAAGGCAGGCTTGGGCTGGGCCAGGACACGTGGG
GTGCCTGGGACGCTGCTTGCCATGCACAGTGATCAGAGAGAGGCTGGGGTGTGTCTGTCCG
GGACCCCCCTGCCTTCCTGCTCACCTACTCTGACCTCCTTCACGTGCCAGGCCTGTGGG
TAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCTACTCTGACCTCCTTCACGTGCCC
AGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 12

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531

><subunit 1 of 1, 327 aa, 1 stop

><MW: 37406, pI: 9.30, NX(S/T): 1

MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHFLFVACLSLGFFSLLWLQLSCSGDVARAVR
 GQGQETSGPPRACPPPEPPPEHWEEDASWGPHRLAVLVPPFRERFEELLVFVPHMRRFLSRKKI
 RHHIYVLNQVDHFRFNRAALINVGFLSSNSTDYIAMHDVDLLPLNEELDYGFPEAGPFHVA
 SPELHPLYHYKTYVGGILLISKQHYRLCNGMSNRFWGWGREDDFYRRIKGAGLQLFRPSGI
 TTGYKTFRHLHDPAWRKRDKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVL
 NIMLDCDKTATPWCTFS

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 29-49 (type II)

N-glycosylation site.

amino acids 154-158

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 27-31

Tyrosine kinase phosphorylation site.

amino acids 226-233

N-myristoylation site.

amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

FIGURE 13

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCT
GCTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCT
CTTCTGCCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTC
CTGCCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGT
GGGGGATGGC**TAA**GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGG
GCTGCCTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGC
CCTGTTCAATTAAGAAATTGTTCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGT
AAATATGTCTTTATAATAAACAGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

p.d.p.c.p. = G.E.S.D.F.

MLLLTLLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFPPSFLCLLPHRPAMTCSQAQ
PRGEGEKVGDG

amino acids 1-15

amino acids 3-18

[illegible]

FIGURE 15

GGGACCCATGCGGCCGTGACCCCCGGCTCCCTAGAGGCCAGCGCAGCCGCGAGCGGACAAAG
 GAGCATGTCCGCGCCGGGAAGGCCCGTCTCCGGCCGCCATAAGGCTCCGGTCCGCGCTGG
 GCGCGCGCCGCGCTCCTGCCCCCGGGCTCCGGGGCGGCCCGCTAGGCCAGTGCGCCGCCG
 CTCGCCCCGCAGGCCCGGGCCGCGAG**CATG**GAGCCACCCGGACGCCGGCGGGGGCCGCGCGCA
 GCCGCCGCTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGAGGCGGCGGCG
 GCGGCGGCGCCGCGGCGCTGCCCGCCGGCTGCAAGCACGATGGGCGGCCCGAGGGGCTGGC
 AGGGCGGCGGGCGCCGCCGAGGGCAAGGTGGTGTGCAGCAGCCTGGAACCTCGCGCAGGTCTCT
 GCCCCCAGATACTCTGCCCAACCGCACGGTCACCCTGATTCTGAGTAACAATAAGATATCCG
 AGCTGAAGAATGGCTCATTTTTCTGGGTAAAGTCTCCTTGAAAGATTGGACCTCCGAAACAAT
 CTTATTAGTAGTATAGATCCAGGTGCCTTCTGGGGACTGTCATCTCTAAAAAGATTGGATCT
 GACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTGAGGACTCACCAATCTGGTTC
 GGCTAAACCTTTCGGGGAATTTGTTTTCTTCATTATCTCAAGGAACCTTTGATTATCTTGCG
 TCATTACGGTCTTTGGAATTCAGACTGAGTATCTTTTGTGTGACTGTAACATACTGTGGAT
 GCATCGCTGGGTAAAGGAGAAGAACATCACGGTACGGGATACCAGGTGTGTTTATCCTAAGT
 CACTGCAGGCCCAACCAGTCACAGGCGTGAAGCAGGAGCTGTTGACATGCGACCCTCCGCTT
 GAATTGCCGTCTTTCTACATGACTCCATCTCATCGCCAAGTTGTGTTTGAAGGAGACAGCCT
 TCCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTGTTGTGGTATCAGGATG
 GGAGAATAGTTGAAACCGATGAATCGCAAGGTATTTTTGTTGAAAAGAACATGATTCACAAC
 TGCTCCTTGATTGCAAGTGCCCTAACCATTCTAATATTCAGGCTGGATCTACTGGAAATTG
 GGGCTGTCATGTCCAGACCAACGTGGGAATAATACGAGGACTGTGGATATTGTGGTATTAG
 AGAGTTCTGCACAGTACTGTCTCCAGAGAGGGTGGTAAACAACAAAGGTGACTTCAGATGG
 CCCAGAACATTGGCAGGCATTACTGCATATCTGCAGTGTACGCGGAACACCCATGGCAGTGG
 GATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGGCGCAGATGTGATAGAGGTGGCT
 TTTGGGCAGATGATGATTATTCTCGCTGTCAGTATGCAATGATGTCACTAGAGTTCTTTAT
 ATGTTTAATCAGATGCCCCCTCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGC
 TTACACTGTGGAAGCAGCCAACCTTTCTGACAAAATGGATGTTATATTTGTGGCAGAAATGA
 TTGAAAAATTTGGAAGATTTACCAAGGAGGAAAAATCAAAGAGCTAGGTGACGTGATGGTT
 GACATTGCAAGTAACATCATGTTGGCTGATGAACGTGTCTGTGGCTGGCGCAGAGGGAAGC
 TAAAGCCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGGCTAGCCGGTG
 GAGCTCACGTTTATTCAACATATTACCCCAATATTGCTCTGGAAGCTTATGTCATCAAGTCT
 ACTGGCTTCACGGGGATGACCTGTACCGTGTTCAGAAAGTGGCAGCCTCTGATCGTACAGG
 ACTTTCGGATTATGGGAGGCGGGATCCAGAGGGGAAACCTGGATAAGCAGCTGAGCTTTAAGT
 GCAATGTTTCAAATACATTTTCGAGTCTGGCACTAAAGGTATGTTACATTCTGCAATCATTT
 AAGACTATTTACAGT**TAA**ATTAGAATGCTCCAAATGTTCTGCTTCGCAAAATAACCTTATTA
 AAAGATTTTTTTTTTGCAAGGAAGATAGGTATTATTGCTTTTGCTACTGTTTTAAAGAAACTA
 ACCAGGAAGAACTGCATTACGACTTTCAAGGGCCCTAGGCATTTTTGCCTTTGATTCCCTTT
 CTTACATAAAAATATCAGAAATTACATTTTATAACTGCAGTGGTATAAATGCAAATATACT
 ATTGTTACATGTGAAAAAATTTTATTTGACTTAAAAGTTTATTTATTTGTTTTTTTGCTCCT
 GATTTTAAGACAATAAGATGTTTTTCATGGGCCCTAAAAGTATCATGAGCCTTTGGCACTGC
 GCCTGCCAAGCCTAGTGGAGAAGTCAACCCTGAGACCAGGTGTTTAATCAAGCAAGCTGTAT
 ATCAAAATTTTTGGCAGAAAACACAAATATGTCATATATCTTTTTTTTAAAAAAGTATTTCA
 TTGAAGCAAGCAAAATGAAAGCATTTTTACTGATTTTTAAAATTGGTGCTTTAGATATATTT
 GACTACACTGTATTGAAGCAATAGAGGAGGCACAACTCCAGCACCCCTAATGGAACCACATT
 TTTTCACTTAGCTTTCTGTGGGCATGTGTAATTGTATTCTCTGCGGTTTTTAATCTCACAG
 TACTTTATTTCTGTCTTGTCCCTCAATAATATCACAAACAATATTCCAGTCATTTTAATGGC
 TGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTAGTGTGAGCACTCAATAAATA
 TTGAATGAATGAACGAAAAAAAAAAAAAAAAA

FIGURE 16

MEPPGRRRRGRAQPPLLLPLSLLALLALLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGK
 VVCSSLELAQVLPPDTLPNRTVTLILSNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGA
 FWGLSSLKRLDLTNNRIGCLNADIFRGLTNLVRNLNSGNLFSSLSQGTFDYLASLRSLEFQT
 EYLLCDCNILWMHRWVKEKNITVRDTRCVYPKSLQAQPVTGVKQELLTCDPPELPSFYMTP
 SHRQVVFEGDSLPPFQCMASYIDQDMQVLWYQDGRIVETDESQGIFVEKNMIHNCSLIASALT
 ISNIQAGSTGNWGCHVQTKRGNNTRTVDIVVLESSAQYCPPERVVNNKGDFRWPRTLGITAY
 YLQCTRNTHGSGIYPGNPQDERKAWRRRCRGGFWADDDYSRCQYANDVTRVLYMFNQMPNLN
 TNAVATARQLLAYTVEAANFSDKMDVIFVAEMIEKFGRFTKEEKSKELGDMVDIASNIMLA
 DERVLWLAQREAKACSRIVQCLQRIATYRLAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCT
 VFQKVAASDRTGLSDYGRRDPEGNLDKQLSFKCNVSNTFSSLALKVCYILQSFKTIYS

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 13-40 (type II)

N-glycosylation site.

amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336,
 433-437, 453-457, 592-596

N-myristoylation site.

amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57,
 57-63, 99-105, 123-129, 142-148, 162-168, 317-323, 320-326,
 384-390, 403-409, 554-560

FIGURE 17

GCGTGGGGATGTCTAGGAGCTCGAAGGTGGTGCTGGGCCTCTCGGTGCTGCTGACGGCGGCC
ACAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTAT
CAGAGACATTGAGAGGCAAATTCGGAAAAAAGAAAACATTCGTCTTTTGGGAGAACAGATTA
TTTTGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAGGATCTCAAAA
TCATTGACTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTGTTGAT
GGAGAGTAGCTTAGTAGTATCTTCATCTTTTTTTTTGGTCACTGTCCTTTTAAACTTGATCA
AATAAAGGACAGTGGGTCATATAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGA
GTGTGGGCAGACACTTTTTGGAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGG
TCACTGTCCAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGA
TGGATAGTAGCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACA
TCGACCTCAGAAGGACTGGAGGAAGGTGAAGTGAGGGAGAGACGCTCCTGATCGTCGAATCC

FIGURE 18

MSRSSKVVVLGLSVLLTAATVAGVHVKKQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILT
EQLEAEREKMLLAKGSQKS

Signal peptide:

amino acids 1-21

18/249

FIGURE 19

CTGTCGTCTTTGCTTCAGCCGCAGTCGCCACTGGCTGCCTGAGGTGCTCTTACAGCCTGTTC
 CAAGTGTGGCTTAATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCCCTCTGCTAAGACC
 GCTGCCATGCCAGTGACGGTAACCCGCACCACCATCACAACCACCACGACGTCATCTTCGGG
 CCTGGGGTCCCCCATGATCGTGGGGTCCCCCTCGGGCCCTGACACAGCCCCCTGGGTCTCCTTCGC
 CTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGTGGGCGCCTGGAC
 GGGGTCCATGGGCAACTGGTCCATGTTACCTGGTGCTTCTGCTTCTCCGTGACCCTGATCA
 TCCTCATCGTGGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGGCGCAACTTCCCC
 ATCACCTTCGCCTGCTATGCGGCCCTCTTCTGCCTCTCGGCCTCCATCATCTACCCCACCAC
 CTATGTCCAGTTCCCTGTCCCACGGCCGTTTCGCGGGACCACGCCATCGCCGCCACCTTCTTCT
 CCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTGGACCCGGGCCCCGGCCCCGGCGAG
 ATCACTGGCTATATGGCCACCGTACCCGGGCTGCTGAAGGTGCTGGAGACCTTCGTTGCCTG
 CATCATCTTCGCGTTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGCCCTGGAGTGGT
 GCGTGGCGGTGTACGCCATCTGCTTCATCCTAGCGGCCATCGCCATCCTGCTGAACCTGGGG
 GAGTGCACCAACGTGCTACCCATCCCCTTCCCCAGCTTCTGTGCGGGGCTGGCCTTGCTGTC
 TGTCCCTCCTCTATGCCACCGCCCTTGTTCCTCTGGCCCCCTCTACCAGTTCGATGAGAAGTATG
 GCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGCAGCCATGCCTACTACGTGTGT
 GCCTGGGACCGCCGACTGGCTGTGGCCATCCTGACGGCCATCAACCTACTGGCGTATGTGGC
 TGACCTGGTGCACCTCTGCCCACCTGGTTTTTTGTCAAGGTCTTAAGACTCTCCCAAGAGGCTCC
 CGTTCCCTCTCCAACCTCTTTGTTCTTCTTGCCCGAGTTTTCTTTATGGAGTACTTCTTTCC
 TCCGCCCTTTCCTCTGTTTTCTCTTCTCCTGTCTCCCCTCCCTCCCACCTTTTTCTTTCTTCC
 CAATTCCCTTGCACTCTAACCAGTTCTTGGATGCATCTTCTTCCCTTTCCCTCTTGCTGT
 TTCCTTCCCTGTGTTGTTTTGTTGCCACATCCTGTTTTTACCCCTGAGCTGTTTCTCTTTTT
 CTTTTCTTTCTTTTTTTTTTTTTTTTTTAAAGACGGATTCTCACTCTGTGGCCAGGCTGGAG
 TGCAGTGGTGCGATCTCAGCTCACTGCAACCCCCGCCCTCCTGGGTTCAAGCGATTCTCCTCC
 CCCAGCCTCCCAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGCACCCAGCCTGTTTCTCTTT
 TTCCACTCTTCTTTTTTCTCATCTCTTTTCTGGGTTGCCTGTCGGCTTTCTTATCTGCCTGT
 TTTGCAAGCACCTTCTCCTGTGTCTTGGGAGCCCTGAGACTTCTTTCTCTCCTTGCCTCCA
 CCCACCTCCAAAGGTGCTGAGCTCACATCCACACCCCTTGACGCCGTCCATGCCACAGCCCC
 CCAAGGGGCCCCATTGCCAAAGCATGCCTGCCACCCCTCGCTGTGCCTTAGTCAGTGTGTAC
 GTGTGTGTGTGTGTGTGTTGGGGGGTGGGGGGTGGGTAGCTGGGGATTGGGCCCCCTTTTCT
 CCCAGTGGAGGAAGGTGTGAGTGTACTTCCCCTTTAAATTAAAAACATATATATATATAT
 ATTTGGAGGTGAGTAATTTCCAATGGGCGGGAGGCATTAAGCACCGACCCTGGGTCCCTAGG
 CCCCCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAGAATTTTTGCCAGGCTTACAGAACAC
 CCACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGCTGGATGCCTTTCATCCCAACTATTCT
 CTGTGGTATGAAAAAG

FIGURE 20

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727

<subunit 1 of 1, 322 aa, 1 stop

<MW: 35274, pI: 8.57, NX(S/T): 1

MPVTVTRTTITTTTTSSSGLGSPMIVGSPRALTQPLGLLRLLQLVSTCVAFSLVASVGAWTG
SMGNWSMFTWCFCFSVTLLIILIVELCGLQARFPLSWRNFPI TFACYAALFCLSASIIYPTTY
VQFLSHGRSRDHAIAATFFSCIACVAYATEVAWTRARPGEITGYMATVPGLLKVLET FVACI
IFAFISDPNLYQHQP ALEWCVAVYAICFILAAIAILLNLGECTNVLP I PFPSFLSGLALLSV
LLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRRLAVAILTAINLLAYVAD
LVHSAHLV FVKV

Important features:**Transmembrane domains:**

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192,
205-226, 235-255 and 294-312

N-glycosylation site.

amino acids 66-69

Glycosaminoglycan attachment site.

amino acids 18-21

FIGURE 21

GAACGTGCCACCATGCCCAGCTAATTTTTGTATTTTATAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTC
TTGAACCTCGTGACCTCATGATCCGCTCACCTCGGCCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGACGC
CTGGCCAGCCTATGCATTTTAAAGAAATTATTCTGTATTAGGTGCTGTGCTAAACATTGGGCCTACAGTGACCA
AAACAGACTGAATTTCCCAAGAGCCAAAGACCAGTGAGGGAGACCAACAAGAAACAGGAAATGCAAAAGAGACCA
TTATTACTACTATGACTAAGGGTCACAAATGGGGTACGTTGATGGAGAGTGATTTGTTAAGAGACTACAGAGGG
AGGACAGACTACCAAGAGGGGGGGCCAGGAAAGCTCCTCTGACGAGGTGGTATTTTCAGCCCAAACCTGGAAGAATGA
GAAAGAGCTAGCCAGCCATCAGAATAGTCCAGAAGAGATGGGGAGCACTACACTCACTACACTTTGGCCTGAGAA
AATAGCATGGGATTGGAGGAGGCTGGGGGAACACCCTTCTGCCGACCTGGGCAGGAGGCATTGAGGGCTTGAGA
AAGGGCAATGGCAGTAGCAGTAGAAAGGACAGGGTAGGAGCAGGGACTTTGCAGGTGGAATCATTAGGTCTTATC
AACAGATATGGGCAAGCAAAGCCAGGGGAGAATTGATGGTAAATGCTGAGGTTTGGAGCCAGGCTAGATGGGACAG
TGGTGGGTGATGCAAAGGAAAGAGGTGAGGAAGCAGGGCCAGACGTGGGGAGAAGGTGTGGGGGTTTGGTTTCCA
TCTTGCCGAGTCTGCCGAATGTGGATGGGAAGACCAAGAGGAGGAGCAAGGGGCAGAGGGGAAGGAATCTTAA
AGAAGTCCCTGGATGCCACACTCTTCTCCTCCTCCTCCTCCTCCTCAGAGGTCTCACTCGTGGTTCTTCAT
TTCCTGCCCTGCCTCCATCTCCTCTGGGTGCTGGGAAAGTGGAGGATTAGCTGAAGTTTGGCTTCTCGGGGCTG
CTGAATCTCCATTGCTTTCTGGGAGGACATAATCACCTGTCTAGCTTCTTATCATCTTACATTCCCTGTAG
CCACTGGGACATATGTGGTGTTCCTTCCTAGCTCCTGTCTCCTCCTCATGCCTTTGCTGGGTATGGGCATGTTAG
GGGGAAGGTCAATTGCTGTGAGAGGGGCACTGACTTTCTAATGGTGTACCCAAGGTGAATGTTGGAGACACAGTC
GCGATGCTGCCAAGTCCCGGCGAGCCCTAACTATCCAGGAGATCGCTGCGCTGGCCAGGTCTCCTCGCATGGT
ATGCAGCCCTCCCAATGTTTCTGGCCACTTTGTCCTTTCTCCTCCCGTTTGCACATCCCTTTGGAACCTGTTTCT
GTGAGTACATGCTGGGGTCTCCCCCTTTCTTCCCTTGGCTCAGGTGAATCTCAGCCCTTCTCCCAACCAAGGTTT
ACATGGATCCTAACTACTGCCACCTTCCACCTCCCTGCACCTGTGCTCCCTGGCCTGGTCTTTACAGGCTTC
TCCACCTCCCTATCTCCAGGTATTTCCAGGTGGTGAAGGACACGTGACCAAGCCTACCGCATGGCCAGG
GCCGAGTGGCTCACCTCATTGAGTGGAAAGGGCTGGAGCAAGCCGAGTGACTCACCTGCTGCCCTGGAATCAGCCT
TTCTCCTATTACAGACCTCAGCGAGGGCGAACAAGAGGCTCGCTTTGCAGCAGGAGTGCTGAGCAGTTTGCCA
TCGCGGAAGCCAAGCTCCGAGCATGGTCTTCCGTTGGATGGCGAGGACTCCACTGATGACTCCTATGATGAGGACT
TTGCTGGGGGAATGGACACAGACATGGCTGGGCAGCTGCCCTGGGGCCGACCTCCAGGACCTGTTACCGGGCC
ACCGTTCTCCCGGCTGTGCGCCAGGGCTCCGTGGAGCCTGAGAGCGACTGCTCACAGACCGTGTCCCGAGACA
CCCTGTGCTCTAGTCTGTGAGCCTGGAGGATGGGTGTTGGGCTCCCGGGCCGGCTGGCCTCCAGCTGCTGG
GCGATGAGCTGCTTCTCGCCAAACTGCCCCCAGCCGGGAAAGTGCCTTCCGCGACCTGGGCCACTGGAGGCCC
AGGACTCACTCTACAACTCGCCCCCTCACAGAGTCTGCCTTTCCCGCGGAGGAGGAGCCAGCCCCCTGCAAGG
ACTGCCAGCCACTCTGCCACCCTAACGGGCAGCTGGGAACGGCAGCGGCAAGCCTCTGACCTGGCCTCTTCTG
GGTGCTGCTCCTTAGATGAGGATGAGGCAGAGCCAGAGGAACAGTGACCCACATCATGCCTGGCAGTGGCATGCA
TCCCCCGGCTGCTGCCAGGGGCGAGGCCTCTGTGCCCAAGTGTGGGCTCAAGGCTCCAGCAGAGCTCCACAGCC
TAGAGGGCTCCTGGGAGCGCTCGCTTCTCCGTTGTGTGTTTTGCATGAAAGTGTTTGGAGAGGAGGCAGGGGCTG
GGCTGGGGGCGCATGTCTGCCCCCACTCCCGGGGCTTGGCCGGGGGTTGCCCGGGGCTCTGGGGCATGGCTACA
GCTGTGGCAGACAGTGATGTTTCTTAAATGCCACACACATTTCTCCTCGGATAATGTGAACCACTA
AGGGGGTTGTGACTGGGCTGTGTGAGGTGGGGTGGGAGGGGGCCAGCAACCCCCACCTCCCCATGCCTCTC
TCTTCTCTGCTTTTCTTCTCACTTCCGAGTCCATGTGAGTGCTTGAAGAATCACCCCCACCTGGAGGGGCTGG
CTCCTGCCCTCCCGGAGCCTATGGGTTGAGCCGTCCCTCAAGGGCCCTGCCAGCTGGGCTCGTGCTGTGCTTC
ATTCACTCTCCATCGTCTCTAAATCTTCTCTTTTTTCTAAAGACAGAAGGTTTTTGGTCTGTTTTTTCAGTC
GGATCTTCTCTTCTCTGGGAGGCTTTGGAATGATGAAAGCATGTACCCTCCACCTTTTCTGGCCCCCTAATGG
GGCCTGGGCCCTTTCCCAACCCCTCCTAGGATGTGCGGGCAGTGCTGGCGCCTCACAGCCAGCCGGGCTGCC
ATTACGCGAGAGCTCTCTGAGCGGAGGTGGAAGAAAGGATGGCTCTGGTTGCCACAGAGCTGGGACTTCATGTT
CTTCTAGAGAGGGCCACAAGAGGGGCCACAGGGGTGGCCGGGAGTTGTCAGCTGATGCCTGCTGAGAGGCAGGAAT
TGTGCCAGTGAGTGACAGTCATGAGGGAGTGCTCTTCTTGGGAGGAAAGAGGTAGAGCCTTTCTGTCTGAAT
GAAAGGCCAAGGCTACAGTACAGGGCCCCGCCAGCCAGGGTGTTAATGCCACGTAGTGGAGGCCTCTGGCAG
ATCCTGCATTCCAAGGTCACTGGACTGTACGTTTTTATGGTTGTGGGAAGGGTGGGTGGCTTTAGAATTAAGGGC
CTTGTAGGCTTTGGCAGGTAAGAGGGCCCAAGGTAAGAACGAGAGCCAACGGGCACAAGCATTTATATATAAGT
GGCTCATAGGTGTTTATTTTGTCTATTTAAGAATTTGTTTTATTAAATTAATATAAAATCTTTGTAAATCTC
TAAAA

FIGURE 22

MFLATLSFLLPFAHPFGTVSCEYMLGSPLSSLAQVNLSPFSSHVKVHMDPNYCHPSTSLHLCS
 LAWSFTRLLHPPLSPGISQVVKDHVTKPTAMAQGRVAHLIEWKGWSKPSDSPAALESASFSSY
 SDLSEGEQEAREFAAGVAEQFAIAEAKLRAWSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPLG
 PHLQDLFTGHRFSRPVRQGSVEPESDCSQTVSPDTLCSSLCSLEDGLLGSPARLASQLLGDE
 LLLAKLPPSRESAFRSLGPLEAQDSLNSPLTESCLSPAEEEEPAPCKDCQPLCPPLTGSWER
 QRQASDLASSGVVSLDEDEAEPEEQ

Signal peptide:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232,
 285-289, 324-328

Tyrosine kinase phosphorylation site.

amino acids 44-52

N-myristoylation site.

amino acids 17-23, 26-32, 173-179

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 11-22

FIGURE 23

GGTTCCTGGGCGCTCTGTTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTCCTT
GGCACCTCCTGCTCAGTGGGACATTGTACACTTAACCCATCTGTTTTCTCTAATGCACGA
CAGATTCCTTTCAGACAGGACAACCTGTGATATTTAGTTTCCTGATTGTAAATACCTCCTAAG
CCTGAAGCTTCTGTTACTAGCCATTGTGAGCTTCAGTTTCTTCATCTGCAAAATGGGCATAA
TACAATCTATTCTTGCCACATCAAGGGATTGTTATTCCTTTAAAAAAAACCAATACCAAAG
AAGCCTACA**ATG**TTGGCCTTAGCCAAAATTCTGTTGATTTCAACGTTGTTTTATTCACTTCT
ATCGGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAAGTTT
TTAAACAATGGAAAATAAACCTATTTCTTTGGAAAGTGAAGCAAACCTAAACTCAGATAAA
GAAAATATAACCACCTCAAATCTCAAGGCGAGTCATTCCCCCTCCTTTGAATCTACCCAACAA
CAGCCACGGAATAACAGATTTCTCCAGTAACATCAGCAGAGCATTCTTTGGGCAGTCTAA
AACCCACATCTACCATTTCCACAAGCCCTCCCTTGATCCATAGCTTTGTTTCTAAAGTGCCT
TGGAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACC
TGCTCTGTCTTCAGAAACTTCACTTGGTCTTTGGTCAATGACACCGTGAAAACCTCCTGATA
ACAGTTCCATTACAGTTAGCATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCTTG
ATAGTGGAACCAAGTGGATGGCTTACCACAAACAGTGATAGCTTCACTGGGTTTACCCCTTA
TCAAGAAAAACAACCTCTACAGCCTACCTTAAAATTCACCAATAATTCAAAACTCTTTCCAA
ATACGTCAGATCCCCAAAAAGAAAATAGAAATACAGGAATAGTATTGCGGGCCATTTTAGGT
GCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTTGTGTGGAAAAAGGAAAAC
GGATTCATTTTCCCATCGGCGACTTTATGACGACAGAAATGAACCAGTTCTGCGATTAGACA
ATGCACCGGAACCTTATGATGTGAGTTTTGGGAATTCTAGCTACTACAATCCAACCTTTGAAT
GATTCAGCCATGCCAGAAAGTGAAGAAAATGCACGTGATGGCATTCCCTATGGATGACATACC
TCCACTTCGTACTTCTGTAT**AGA**ACTAACAGCAAAAAGGCGTTAAACAGCAAGTGTCATCTA
CATCCTAGCCTTTTGACAAATTCATCTTTCAAAGGTTACACAAAATTACTGTCACGTGGAT
TTTGTCAGGAGAATCATAAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCAT
CCAAAGGTTTTCTTTCTTACAATTTTTTGCCATCCTGAGGCATTTACTAAGTAGCCTTAATT
TGTATTTTAGTAGTATTTTCTTAGTAGAAAATATTTGTGGAATCAGATAAACTAAAAGATT
TCACCATTACAGCCCTGCCTCATAACTAAATAATAAAAATTATTCCACCAAAAAATTCTAAA
ACAATGAAGATGACTCTTTACTGCTCTGCCTGAAGCCCTAGTACCATAATTCAAGATTGCAT
TTTCTTAAATGAAAATTGAAAGGGTGCTTTTTAAAGAAAATTTGACTTAAAGCTAAAAAGAG
GACATAGCCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTC
TAGTACGTTATAATTTCTAGATCAGCACACACATGATCAGCCCCTGAGTTATGAAGCTGA
CAATGACTGCATTCAACGGGGCCATGGCAGGAAAGCTGACCCTACCCAGGAAAGTAATAGCT
TCTTTAAAAGTCTTCAAAGGTTTTGGGAATTTTAACTTGTCTTAATATATCTTAGGCTTCAA
TTATTTGGGTGCCTTAAAACTCAATGAGAATCATGGT

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732

><subunit 1 of 1, 334 aa, 1 stop

><MW: 36294, pI: 4.98, NX(S/T): 13

MLALAKILLISTLFYSLLSGSHGKENQDINTTQNI AEVFKTMENKPISLESEANLNSDKENI
TTSNLKASHSPPLNLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSFVSKVPWNA
PIADEDLLPISAHPNATPALSSENFTWSLVNDTVKTPDNSSITVVSILSSEPTSPSVTPLIVE
PSGWLTTNSDSFTGFTPYQEKTTLQPTLKFTNNSKLFNPNTSDPQKENRNTGIVFGAILGAIL
GVSLTLVGYLLCGKRKTDSFSHRRLYDDRNEPVLRLDNAPEPYDVSFGNSSYYNPTLND
SA
MPESEENARDGIPMDDIPPLRTSV

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 235-262

N-glycosylation site.

amino acids 30-34, 61-65, 79-83, 90-94, 148-152, 155-159,
163-167, 218-222, 225-229, 298-302, 307-311

FIGURE 25

AACAGGATCTCCTCTTGAGTCTGCAGCCCAGGACGCTGATTCCAGCAGCGCCTTACCGCGC
AGCCCGAAGATTCAC**TATG**GTGAAAATCGCCTTCAATACCCCTACCGCCGTGCAAAAGGAGG
AGGCGCGGCAAGACGTGGAGGCCCTCCTGAGCCGCACGGTCAGAACTCAGATACTGACCGGC
AAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGCTCCTCTGGGAGATGTATGCTTACTCT
CTTAGGCCTTTCATTCATCTTGGCAGGACTTATTGTTGGTGGAGCCTGCATTTACAAGTACT
TCATGCCCAAGAGCACCATTTACCGTGGAGAGATGTGCTTTTTTGATTCTGAGGATCCTGCA
AATTCCCTTCGTGGAGGAGAGCCTAACTTCCTGCCTGTGACTGAGGAGGCTGACATTCGTGA
GGATGACAACATTGCAATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCCTGCAG
CAATTATTCATGACTTTGAAAAGGGAATGACTGCTTACCTGGACTTGTTGCTGGGGAACCTGC
TATCTGATGCCCCCTCAATACTTCTATTGTTATGCCTCCAAAAAATCTGGTAGAGCTCTTTGG
CAA**ACT**GGCGAGTGGCAGATATCTGCCTCAA**ACT**TATGTGGTTCGAGAAGACCTAGTTGCTG
TGGAGGAAATTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGA
AAGTCCTTCCGCCTTCGTGCGAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAA
ATGCTGGAAGATTAGACACTTCCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAG**T**
AAGAGGCAACAGATAGAGTGTCTTGGTAATAAGAAGTCAGAGATTTACAATATGACTTTAA
CATTAAAGGTTTATGGGATACTCAAGATATTTACTCATGCAATTTACTCTATTGCTTATGCTTT
AAAAAAAGGAAAAAAAAAAAAAAAACTACTAACCCTGCAAGCTCTTGTCAAATTTTAGTTTAAT
TGGCATTGCTTGTTTTTTGAAACTGAAATTACATGAGTTTCATTTTTCTTTGCATTTATAG
GGTTTAGATTTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATTCCATCC
GTTGTTTTTTTTGTTTGTTTGTTTTTCTTTTCTTTAAGTAAGCTCTTTATTCATCTTATG
GTGGAGCAATTTTAAATTTGAAATATTTTAAATTGTTTTTGAACTTTTTGTGTAAAATATA
TCAGATCTCAACATTGTTGGTTTCTTTTGTTTTTCATTTTGTACAACCTTCTTGAATTTAGA
AATTACATCTTTGCAGTCTGTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAATT
TTCATGAGACAGTCATTTTTAACTAATGCAGTGATTCTTTCTCACTACTATCTGTATTGTGG
AATGCACAAAATTGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCTACAA
CCCTATAATAAATTTTACTCTATACAAAAAAAAAAAAAAAAAAAA

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828

<subunit 1 of 1, 263 aa, 1 stop

<MW: 29741, pI: 5.74, NX(S/T): 1

MVKIAFNTPTAVQKEEARQDVEALLSRTVVRTQILTGKELRVATQEKEGSSGRCMLTLLGLSF
 ILAGLIVGGACIYKYFMPKSTIYRGEMCFDSEDPANSLRGGEPNFLPVTEEADIREDDNIA
 IIDVPVPSFSDSDPAAIHDFEKGMTAYLDLLLGNCYLMPLNTSIVMPPKNLVELFGKLAGS
 RYLPQTYVVREDLVAVEEIRDVSNLGIFIYQLCNNRKSFRLLRRDLLLGFNKRAIDKCWKIR
 HFPNEFIVETKICQE

Type II transmembrane domain:

amino acids 53-75

N-glycosylation site.

amino acids 166-170

Casein kinase II phosphorylation site.

amino acids 35-39, 132-136, 134-138

N-myristoylation site.

amino acids 66-72, 103-109

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 63-74

FIGURE 27

GGAGGAGGGAGGGCGGGCAGGCGCCAGCCCAGAGCAGCCCCGGGCACCAGCACGGACTCTCT
 CTTCCAGCCCAGGTGCCCCCACTCTCGCTCCATTCGGCGGGAGCACCCAGTCCTGTACGCC
 AAGGAACTGGTCCTGGGGGCACC**ATG**GTTTCGGCGGCAGCCCCAGCCTCCTCATCCTTCTG
 TTGCTGCTGCTGGGGTCTGTGCCTGCTACCGACGCCCCGCTCTGTGCCCCCTGAAGGCCACGTT
 CCTGGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGGCTCGTCGGCCTCCTCCCCGAGCCTCC
 CGCCACCCTGGACCCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCCACAACCCTGGGG
 GGCCCATCACCCCCACCAACTTCCTGGATGGGATAGTGGACTTCTTCCGCCAGTACGTGAT
 GCTGATTGCTGTGGTGGGCTCCCTGGCCTTTCTGCTGATGTTTCATCGTCTGTGCCGCGGTCA
 TCACCCGGCAGAAGCAGAAGGCCTCGGCCTATTACCCATCGTCCTTCCCCAAGAAGAAGTAC
 GTGGACCAGAGTGACCGGGCCGGGGGCCCCGGGCCTTCAGTGAGGTCCCCGACAGAGCCCC
 CGACAGCAGGCCCCGAGGAAGCCCTGGATTCCCTCCCGGCAGCTCCAGGCCGACATCTTGGCCG
 CCACCCAGAACCTCAAGTCCCCCACCAGGGCTGCACTGGGCGGTGGGGACGGAGCCAGGATG
 GTGGAGGGCAGGGGCGCAGAGGAAGAGGAGAAGGGCAGCCAGGAGGGGGACCAGGAAGTCCA
 GGGACATGGGGTCCCAGTGGAGACACCAGAGGCGCAGGAGGAGCCGTGCTCAGGGGTCTTG
 AGGGGGCTGTGGTGGCCGGTGAGGGCCAAGGGGAGCTGGAAGGGTCTCTCTTGTAGCCAG
 GAAGCCCAGGGACCAGTGGGTCCCCCGAAAGCCCCTGTGCTTGCAGCAGTGTCCACCCAG
 GTGCT**TAA**CAGTCCTCCCGGGCTGCCAGCCCTGACTGTGCGGGCCCCCAAGTGGTCACCTCCCC
 GTGTATGAAAAGGCCTTCAGCCCTGACTGCTTCCTGACACTCCCTCCTTGGCCTCCCTGTGG
 TGCCAATCCCAGCATGTGCTGATTCTACAGCAGGCAGAAATGCTGGTCCCCGGTGCCCCGGA
 GGAATCTTACCAAGTGCCATCATCCTTCACCTCAGCAGCCCCAAAGGGCTACATCCTACAGC
 ACAGCTCCCCTGACAAAGTGAGGGAGGGGCACGTGTCCCTGTGACAGCCAGGATAAAACATCC
 CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCCCAAACCTACTTTTTTAAACA
 GCTACAGGGTAAAATCCTGCAGCACCCACTCTGGAAAATACTGCTCTTAATTTTCTGAAGG
 TGGCCCCCTGTTTCTAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTTAATCCTC
 TCAAGCGCTCTCCAAGCACCCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGCTGG
 GATCAGGTTGAATGAATGGAACCTCTCCTGTCTGGCCTCAAAGCAGCCTAGAAGCTGAGGG
 GCTGTGTTTGAGGGGACCTCCACCCTGGGGAAGTCCGAGGGGCTGGGGAAGGGTTTCTGACG
 CCCAGCCTGGAGCAGGGGGGGCCCTGGCCACCCCTGTTGCTCACACATTGTCTGGCAGCCTG
 TGTCCACAATATTCGTCAGTCCTCGACAGGGAGCCTGGGCTCCGTCCTGCTTTAGGGAGGCT
 CTGGCAGGAGGTCTCTCCCCATCCCTCCATCTGGGGCTCCCCAACCTCTGCACAGCTCT
 CCAGGTGCTGAGATATAATGCACCAGCACAATAAACCTTTATTCCGGCCTGAAAAAAAAAAAA
 AAAGA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852

><subunit 1 of 1, 283 aa, 1 stop

><MW: 29191, pI: 4.52, NX(S/T): 0

MVSAAAPSLILLILLLLGSPATDARSVPLKATFLEDVAGSGEAEAGSSASSPSLPPPWTPAL
SPTSMGPQPTTLGGPSPPTNFDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKA
SAYYPSSFPPKKKYVDQSDRAGGPRAFSEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSP
TRAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVPVETPEAQEEPCSGVLEGAVVAGE
GQGELEGSLLLAQEAQGPVGPPEPCACSSVHPSV

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 94-118

N-myristoylation site.

amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199,
211-217, 238-244, 242-248

FIGURE 29

[illegible]

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212

><subunit 1 of 1, 440 aa, 1 stop

><MW: 42208, pI: 6.36, NX(S/T): 1

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
 AGSKVSEALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
 ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGNPGGLGTPWVHGYPGNSAGSFGM
 NPQGAPWGGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
 GSQSGSSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGNSGGSRGDSGSESSW
 GSSTGSSSGNHGGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRL
 GGSGDNRYRGQSSWGSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDINKDQ
 RSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
 74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
 155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
 218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
 249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
 279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
 295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
 323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
 383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 31

GACCGGTCCCTCCGGTCTGGATGTGCGGACTCTGCTGCAGCGAGGGCTGCAGGCCCGCCGGGCGGTGCTCACCG
 TGCCCTGGCTGGTGGAGTTTCTCTCCTTTGCTGACCATGTTGTTCCCTTGCTGGAATATTACCGGGACATCTTCA
 CTCTCCTGCTGCGCCTGCACCGGAGCTTGGTGTGTGCGCAGGAGAGTGAGGGGAAGATGTTTCTTCTGAACAAGC
 TGCTGCTACTTGCTGTCTGGGCTGGCTTTTCCAGATTTCCACAGTCCCTGAGGACTTGTTCTTTCTGGAAGAGG
 GTCCCTCATATGCCCTTGAGGTGGACACAGTAGCCCCAGAGCATGGCTTGGACAATGCGCCTGTGGTGGACCAGC
 AGCTGCTCTACACCTGCTGCCCTACATCGGAGAGCTCCGGAACTGCTCGCTTCGTGGGTGTGAGGCAGTAGTG
 GACGGAGTGGGGGCTTCATGAGGAAAATCACCCCCACCCTACCACCAGCCTGGGAGCCCAGCCTTCCCAGACCA
 GCCAGGGGCTGCAGGCACAGCTCGCCCAGGCCTTTTCCACAACCAGCCGCCCTCCTTGCGCCGGACCGTAGAGT
 TCGTGGCAGAAAGAATTGGATCAAATGTGTCAAACATATCAAGGCTACACTGGTGGCAGATCTGGTGCGCCAGG
 CAGAGTCACTTCTCCAAGAGCAGCTGGTGACACAGGGAGAGGAAGGGGGAGACCCAGCCAGCTGTTGGAGATCT
 TGTGTTCCAGCTGTGCCCTCACGGGGCCAGGCATTGGCCCTGGGGCGGGAGTTCTGTCAAAGGAAGAGCCCTG
 GGGCTGTGCGGGCGCTGCTTCCAGAGGAGACCCGGCAGCCGTTCTGAGCAGTGCAGAGAACATTGCTGTGGGGC
 TTGCAACAGAGAAAGCCTGTGCTTGGCTGTGAGCCAACATCACAGCACTGATCAGGAGGGAGGTGAAAGCAGCAG
 TGAGTCGCACACTTCGAGCCCAGGGTCTGAACCTGCTGCCCGGGGGAGCGGAGGGGCTGCTCCCGCGCCCTGAC
 GTGCTCTCCTTGGCCGTGGGGCCACGGGACCCTGACGAGGGAGTCTCCCCAGAGCATCTGGAACAGCTCCTAGGC
 CAGCTAGAGCCAGAGCTGCGGTGCCGCCAGTTCTGTGCCACCTGCTGAGCAGCATCTGGCAAAGTCTGTG
 GAGTTAGCTTCCCTCCTCGTTGCAGATCAAATTCCTATCCTAGGGCCCCCGGCACAGTACAGGCTGGAGAGAGGG
 CAGGCTCGAAGGCTTCTGCACATGCTGCTTTCTTGTGGAAGGAAGACTTTTCCAGGGGCCGTTCCGCTGCAGCTG
 CTGCTGAGCCCAAGAAATGTGGGCTTCTGGCAGACACAAGGCCAAGGAGTGGGACTTGCTGCTATTCTTGCTA
 CGGGAGCTGGTGGAGAAGGGTCTGATGGGACGGATGGAGATAGAGGCTGCTGGGCGAGCCTCCACCAGGCCCCAG
 TGGCCAGGGGACTTTGCTGAAGAATTAGCAACACTGTCTAATCTGTTTCTAGCCGAGCCCCACCTGCCAGAACCC
 CAGCTAAGAGCCTGTGAGTTGGTGCAGCCAAACCGGGGCACTGTGCTGGCCCCAGAGCTAGGGCTGAGAAGTGGCC
 CTGCCCTTGGGCATTGCACCAGAACCCTGGACCCCGCCTCACGAGGAGGCCCAAGTGCCCAATGCAGACCCCTCAC
 TGGTTGGGGTGTAGCTGGGTCTACAGTCACTTCTGCTCTAAGGGTGTCACTGCCTGGCATCCCACCACGCGA
 ATCCTAGAGGAAGGAGAGTTGGCCTGATTTGGGATTATGGCAGAAAAGTCCAGAGATGCCAGTCTGGAGTAGAA
 GAGGTGGTGTGTTTATCTCTTGGATACTAAATGAAATGAGGTGTGTGGGCTTGTCAACACAGAATTCAAGCCT
 CATTTGCTATCCAGCATCTCTTAAACTTTGTAGTCTTGAATTCATGACAGAGGCAAATGACTCCTGCTTAAC
 TTATGAAGAAAGTTAAACATGAATCTTGGGAGTCTACATTTTCTTATCACCAGGAGCTGGACTGCCATCTCCTT
 ATAAATGCCTAACACAGGCCGGGTCTGGTGGCTCATGCCTGTAATCCCAGCACTTTGAGAGGCCTGAGGTGCGCG
 GACTGCCTGAGGTGAGGAATTCAGAGCCAGCCTGGCCAACATGGCAAAACCCCATCTCTACTAAAAATAAAAAA
 TTATTAGCTGGGCATGGTGGTGTGTGCCTGTAATCCCAGCTACTCAGGAGGATGAGGCAGGAGACCTGCTTGAAC
 CTGGAGGTGGAGGTTGCAGTGAGCCGAGGTGCGACCCTGCACTCCAGTCTGGGTAACAGAGCGAGACTTTCTAG
 AAAAGCCTAACAAACAGATAAGGTAGGACTCAACCACTGAAACCTGACTTTCCCCCTGTACCTTCAGCCCCCTG
 TGCAGGTAGTAACCTCTTGAACCTCTCCCTGACCAGGGACCAAGCACAGGGCATTAGAGCTTTTTTAGAATAAA
 CTGGTTTTCTTTAAAAAAGGGGCTTTTATTAATTTCTCCCCACACGATGGCTCCTGCAATCTGCCACAGCTC
 TGGGGCGTGTCTGTAGGGAAGGCCCTGTTTTCCCTGAGGCGGGGCTGGGCTTGTCCATGGGTCCGCGGAGCTG
 CGCTGTCTTGGCGCCTGGCGTGTGTCTAGCTGCTTCTTGGCGGCACAGAGCTGCGGGGTCTGGGGGCACCGGG
 AGCTAAGAGCAGGCTCTGGTGCAGGGGTGGAGGCCTGTCTCTTAACCGACACCCTGAGGTGCTCCTGAGATGCTG
 GGTCCACCCTGAGTGGCACGGGGAGCAGCTGTGGCCGGTGTCTCTCYTAGGCCAGTCTGGGGAACTAAGCTC
 GGGCCCTTCTTTGCAAAGACCGAGGATGGGGTGGGTGTGGGGGACTCATGGGGAATGGCCTGAGGAGCTACGTGT
 GAAGAGGGCGCCGGTTTGTGGCTGCAGCGGCCTGGAGCGCCTCTCTCCTGAGCCTCAGTTTCCCTTCCGTCTA
 ATGAAGAACATGCCGTCTCGGTGTCTCAGGGCTATTAGGACTTGCCCTCAGGAAGTGGCCTTGGACGAGCGTCAT
 GTTATTTTCACAACTGTCTGCGACGTTGGCCTGGGCACGTCATGGAATGGCCCATGTCCCTCTGCTGCGTGGAC
 GTCGCGGTGCGGAGTGCGCAGCCAGAGGCGGGGCCAGAGCTGCGCCTGGGGGTGAGGGGAGGCGCCCCGGGAGGG
 CCTCACAGGAAGTTGGGCTCCCGCACACCAGGCAGGGCGGGCTCCCGCCGCGCGCCGCCACCACCGTCCAGG
 GGCCGGTAGACAAAGTGGAAGTGCGCCTTGGGCTCGCTGCGCAGCAGGTAGCCCTTGATGCAGTGCGGCAGCGCG
 TCGTCCGCCAGCTGGAAGCAGCGCCCTCCACCAGCACGAACAGCCGGTGCCT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 32

MCFLNKLALLLAVLGWLFQIPTVPEDLFFLEEGPSYAFEVDTVAPEHGLDNAPVVDQQLLYTC
CPYIGELRKLLASWVSGSSGRSGGFMRKITPTTTTSLGAQPSQTSQGLQAQLAQAFFHNQPP
SLRRTVEFVAERIGSNCVKHIKATLVADLVRQAESLLQEQLVTQGEEGGDPAQLLEILCSQL
CPHGAQALALGREFCQRKSPGAVRALLPEETPAAVLSSAENIAVGLATEKACAWLSANITAL
IRREVKA AVSRTLRAQGPEPAARGERRGCSRA

Signal peptide:

amino acids 1-18

N-glycosylation site.

amino acids 244-248

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 89-93

Casein kinase II phosphorylation site.

amino acids 21-25, 167-171, 223-227

N-myristoylation site.

amino acids 100-106, 172-178, 207-213

Microbodies C-terminal targeting signal.

amino acids 278-282

FIGURE 33

TCCCTTGACAGGTCTGGTGGCTGGTTCGGGGTCTACTGAAGGCTGTCTTGATCAGGAACTG
 AAGACTCTCTGCTTTTGCCACAGCAGTTCCTGCAGCTTCCTTGAGGTGTGAACCCACATCCC
 TGCCCCCAGGGCCACCTGCAGGACGCCGACACCTACCCCTCAGCAGACGCCGGAGAGAAATG
 AGTAGCAACAAAGAGCAGCGGTGAGCAGTGTTCGTGATCCTCTTTGCCCTCATCACCATCCT
 CATCCTCTACAGCTCCAACAGTGCCAATGAGGTCTTCCATTACGGCTCCCTGCGGGGCCGTA
 GCCGCCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGC
 AACAAGACACTGCCCTCTCGGTGCCACCAGTGTGTGATTGTCAGCAGCTCCAGCCACCTGCT
 GGGCACCAAGCTGGGCCCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGCAC
 CCACCACTGGCTACTCAGCTGATGTGGGCAACAAGACCACCTACCGCGTCGTGGCCCCATTCC
 AGTGTGTTCCGCGTGTGAGGAGGCCCCAGGAGTTTGTCAACCGGACCCCTGAAACCGTGT
 CATCTTCTGGGGGCCCCCGAGCAAGATGCAGAAGCCCCAGGGCAGCCTCGTGCGTGTGATCC
 AGCGAGCGGGCCTGGTGTTCCTCAACATGGAAGCATATGCCGTCTCTCCCGGCCGATGCGG
 CAATTTGACGACCTCTTCCGGGGTGAGACGGGCAAGGACAGGGAGAAGTCTCATTCGTGGTT
 GAGCACAGGCTGGTTTACCATGGTGTGCGGGTGGAGTTGTGTGACCACGTGCATGTCTATG
 GCATGGTCCCCCCCCAACTACTGCAGCCAGCGGCCCGCCTCCAGCGCATGCCCTACCACTAC
 TACGAGCCCCAAGGGGCCGGACGAATGTGTACCTACATCCAGAATGAGCACAGTCGCAAGGG
 CAACCACCACCGCTTCATCACCAGAGAAAAGGGTCTTCTCATCGTGGGCCCAGCTGTATGGCA
 TCACCTTCTCCCACCCCTCCTGGACCTAGGCCACCCAGCCTGTGGGACCTCAGGAGGGTCAG
 AGGAGAAGCAGCCTCCGCCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTG
 CTGGAGTGTCTCCCAGCCAATCAGGGCCTTGAGGAGGATGTATCCTCCAGCCAATCAGGGCC
 TGGGGAATCTGTTGGCGAATCAGGGATTGGGAGTCTATGTGGTTAATCAGGGGTGTCTTTC
 TTGTGCAGTCAGGGTCTGCGCACAGTCAATCAGGGTAGAGGGGGTATTTCTGAGTCAATCTG
 AGGCTAAGGACATGTCCTTTCCCATGAGGCCTTGGTTCAGAGCCCCAGGAATGGACCCCCCA
 ATCACTCCCCACTCTGCTGGGATAATGGGGTCTGTCCCAAGGAGCTGGGAACCTTGGTGTG
 CCCCCTCAATTTCCAGCACCAGAAAGAGAGATTGTGTGGGGGTAGAAGCTGTCTGGAGGCC
 GGCCAGAGAATTTGTGGGGTGTGGAGGTTGTGGGGGCGGTGGGGAGGTCCCAGAGGTGGGA
 GGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTGGACAAACCCTTCCCCCTCTCTGGG
 CACCCTTCTGCCCACACCAAGTTTCCAGTGCAGGAGTCTGAGACCCTTTCCACCTCCCCTACAA
 GTGCCCTCGGGTCTGTCTCTCCCGTCTGGACCTCCAGCCACTATCCCTTGCTGGAAGGCT
 CAGCTCTTTGGGGGGTCTGGGGTGACCTCCCCACCTCCTGGAAAACCTTTAGGGTATTTTTGC
 GCAAACCTCCTTCAGGGTTGGGGGACTCTGAAGGAAACGGGACAAAACCTTAAGCTGTTTTCT
 TAGCCCCCTCAGCCAGCTGCCATTAGCTTGGCTCTTAAAGGGCCAGGCCTCCTTTTCTGCCCT
 CTAGCAGGGAGGTTTTCCAACGTGTTGGAGGCGCCTTTGGGGCTGCCCTTTGTCTGGAGTCA
 CTGGGGGCTTCCGAGGGTCTCCCTCGACCCTCTGTCTCCTGGGATGGCTGTCTGGGAGCTGT
 ATCACCTGGGTTCTGTCCCCTGGCTCTGTATCAGGCACTTTATTAAAGCTGGGCCTCAGTGG
 GGTGTGTTTGTCTCCTGCTCTTCTGGAGCCTGGAAGGAAAGGGCTTCAGGAGGAGGCTGTGA
 GGCTGGAGGGACCATGGAGGAGGCCAGCAGCTAGCCATTGCACACTGGGGTGATGGGTGG
 GGGCGGTGACTGCCCCAGACTTGGTTTTGTAAATGATTTGTACAGGAATAAACACACCTACGC
 TCCGGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 34

MSSNKEQRSADFVILFALITILILYSSNSANEVFHYGSLRGRSRRPVNLKKWSITDGYVPIL
 GNKTLPSRCHQCVIVSSSSHLLGTKLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAH
 SSVFRVLRRPQEFVNRTPETVFIFWGPPSKMQKPQGSLVRVIQRAGLVFPNMEAYAVSPGRM
 RQFDDLFRGETGKDREKSHSWLSTGWFTMVIARELCHVHVYGMVPPNYCSQRPRQLQRMPIH
 YYEPKGPDECVTYIQNEHSRKGNNHRRFITEKRVFSSWAQLYGITFSHPSWT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 9-31 (type II)

N-glycosylation site.

amino acids 64-68, 115-119

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 50-54

Casein kinase II phosphorylation site.

amino acids 3-7, 29-33, 53-57, 197-201

Tyrosine kinase phosphorylation site.

amino acids 253-262

N-myristoylation site.

amino acids 37-43, 114-120, 290-294

FIGURE 35

GTTTCATAGTTGGCGTCTTCTAAAGGAAAAACACTAAAATGAGGAACCTCAGCGGACCGGGAGCGACGCAGCTT
 GAGGGAAGCATCCCTAGCTGTTGGCGCAGAGGGGCGAGGCTGAAGCCGAGTGGCCCCGAGGTGTCTGAGGGGCTGG
 GGCAAAGGTGAAAGAGTTTTCAGAACAAAGCTTCTTGGAAACCCATGACCCATGAAGTCTTGTGACATTTATACCGT
 CTGAGGGTAGCAGCTCGAACTAGAAGAAGTGGAGTGTGGCAGGGACGGCAGTATCTCTTTGTGTGACCCTGGC
 GGCTTATGGGACGTTGGCTTCAGACCTTTGTGATACACCATGTGCGTGGGACGATGACGGCGTGGAGAGGAATG
 AGGCTTGAAGTACACTGGCTTGCCTCCTCTAGCCACAGCAGGCTGCTTTGCTGACTTGAACGAGGTCCCTCAG
 GTACCCGTCCAGCTGCGTCCACCCGTCAGAAAGCCGAGGACACTGTGATCTTGGCTGCGTGGTGGAAACCTCCA
 AGGATGAATGTAACCTGGCGCTGAATGGAAAGGAGCTGAATGGCTCGGATGATGCTCTGGGTGTCTCATCACC
 CACGGGACCTCGTCATCACTGCCCTTAAACAACCACACTGTGGGACGGTACCAGTGTGTGGCCCGGATGCCTGCG
 GGGCTGTGGCCAGCGTGCCAGCCACTGTGACACTAGCCAATCTCCAGGACTTCAAGTTAGATGTGCAGCACGTG
 ATTGAAGTGGATGAGGAAACACAGCAGTCAATTGCTGCCACTGCTGAGAGCCACCCCAAAGCCAGGTCCGG
 TACAGCGTCAAACAAGAGTGGCTGGAGGCCCTCCAGAGGTAACCTGATCATGCCCTCAGGGAACCTCCAGATT
 GTGAATGCCAGCCAGGAGGACGAGGGCATGTACAAGTGTGCAGCCTACAACCCAGTGACCCAGGAAGTGAAAACC
 TCCGGCTCCAGCGACAGGCTACGTGTGCGCCGCTCCACCCGCTGAGGCTGCCCGCATCATCTACCCCCAGAGGCC
 CAAACCATCATCGTCACCAAAGCCAGAGTCTATTCTGGAGTGTGTGGCCAGTGAATCCACCCCCAGCGGT
 ACCTGGGCCAAGGATGGGTCCAGTGTACCCGGCTACAACAAGACGCGCTTCTGCTGAGCAACCTCCTCATCGAC
 ACCACCAGCGAGGAGGACTCAGGCACCTACCGCTGCATGGCCGACAATGGGGTGGGCGAGCCCGGGGACGCGGT
 ATCCTCTACAATGTCCAGGTGTTTGAACCCCTGAGGTACCATGGAGCTATCCAGCTGGTTCATCCCCTGGGG
 CAGAGTGCCAAAGCTTACCTGTGAGGTGCGTGGGAACCCCGCCCTCCGTGCTGTGGCTGAGGAATGCTGTGCC
 CTCTATCTCCAGCCAGCGCTCCGGCTCTCCGCGAGGGCCCTGCGCGTGTCTCAGCATGGGGCTGAGGACGAAGGC
 GTCTACCAAGTGCATGGCCGAGAAGCAGGTTGGGAGCGCCCATGCCGTAGTCCAGCTGCGGACCTCCAGGCCAAGC
 ATAACCCCAAGGCTATGGCAGGATGCTGAGTGGCTACTGGCACACCTCCTGTATCACCTCCAACTCGGCAAC
 CCTGAGCAGATGCTGAGGGGGCAACCGGCCCTCCCCAGACCCCAACGTCAGTGGGGCCTGCTTCCCCGAAGTGT
 CCAGGAGAGAAGGGGACGGGGGCTCCCGCGAGGCTCCCATCATCTCAGCTCGCCCCGACCTCCAAGACAGAC
 TCATATGAAGTGGTGTGGCGGCCCTCGGCATGAGGGCAGTGGCCGGGCGCCAACTCTACTATGTGGTGAAACAC
 CGCAAGCAGGTCAAAATTCCTCTGACGATTGGACATCTCTGGCATTCAGCCAACAGCAGCCGCTGACCTC
 ACCAGACTTGACCCCGGGAGCTTGTATGAAGTGGAGATGGCAGCTTACAACCTGTGCGGGAGAGGGCCAGACGCC
 ATGGTCACCTTCCGAAGTGGACGGCGGCCAAACCCGAGATCATGGCCAGCAAAGAGCAGCAGATCCAGAGAGAC
 GACCTGGAGCCAGTCCCCAGAGCAGCAGCCAGCCAGACCAGGCCGCTCTCCCCCAGAAAGCTCCCGACAGG
 CCCACATCTCCACGGCTCCGAGACCTCAGTGACGTGACCTGGATTCCCGTGGGAATGGTGGGTCCCAATC
 CAGTCTTCCGTGTGGAGTACAAGAAGCTAAAGAAAGTGGGAGACTGGATTCTGGCCACCAGCGCCATCCCCCA
 TCGCGGCTGTCCGTGGAGATCACGGCCCTAGAGAAAGGCACCTCCTACAAGTTTCGAGTCCGGGCTCTGAACATG
 CTGGGGGAGAGCGAGCCAGCGCCCCCTCTCGGCCCTACGTGGTGTGCGGGCTACAGCGGTGCGGTGTACGAGAGG
 CCGTGGCAGGTCTTATATACCTTACGGATGCGGTCAATGAGACCAACCATCATGCTCAAGTGGATGTACATC
 CCAGCAAGTAACAACAACACCCCAATCCATGGCTTTTATATCTATTATCGACCCACAGACAGTGAACATGATAGT
 GACTACAAGAAGGATATGGTGGAAAGGGGACAAGTACTGGCACTCCATCAGCCACCTGCAGCCAGAGACCTCCTAC
 GACATTAAGATGCAGTGCTTCAATGAAGGAGGGGAGAGCGAGTTGAGCAACGTGATGATCTGTGAGACCAAGCT
 CGGAAGTCTTCTGGCCAGCCTGGTGCAGTGCACCCCAACTCTGGCCCCACACAGCCGCCCTTCTCTGAAACC
 ATAGAGCGGCCGGTGGGCACTGGGGCCATGGTGGCTCGCTCCAGCGACCTGCCCTATCTGATTGTGCGGGTCTGTC
 CTGGGCTCCATCGTTCTCATCATCGTCACTTCACTCCCTTCTGCTTGTGGAGGGCCTGGTCTAAGCAAAAACAT
 ACAACAGACCTGGGTTTTCTCGAAGTGCCTTCCACCTCCTGCCGTATACTATGGTGCCATTGGGAGGACTC
 CCAGGCCACCAGGCCAGTGGACAGCCCTACCTCAGTGGCATCAGTGGACGGGCCTGTGCTAATGGGATCCACATG
 AATAGGGGCTGCCCCCTCGGCTGCAGTGGGCTACCGGGCATGAAGCCCCAGCAGCACTGCCAGGCGAGCTTCAG
 CAGCAGAGTGACACCAGCAGCCTGCTGAGGCGAGACCCATCTTGGCAATGGATATGACCCCCAAAGTACCAGATC
 ACGAGGGGTCCCAAGTCTAGCCCGACGAGGGCTCTTTCTTATACACACTGCCCGACGACTCCACTCACCAGCTG
 CTGCAGCCCCATCACGACTGCTGCCAACGCCAGGAGCAGCCTGCTGCTGTGGGCCAGTCAGGGGTGAGGAGAGCC
 CCGACAGTCTGTCTTGAAGCAGTGTGGGACCTCCATTTCACTCAGGGCCCCCATGCTGCTTGGGCCCTTGTG
 CCAGTTGAAGAGGTGGACAGTCTGACTCCTGCCAAGTGAAGTGGAGGAGACTGGTGTCCCCAGCACCCGTAGGG
 GCCTACGTAGGACAGGAACCTGGAATGCAGCTCTCCCGGGGCCACTGGTGGCTGTGTCTTTTGAACACCACT
 CTCACAATTTAGGCAGAAGCTGATATCCCAGAAAGACTATATATTGTTTTTTTTTAAAAAAAAGAGAAAAA
 AGAGACAGAGAAAATTGGTATTTATTTTCTATTATAGCCATATTTATATATTTATGCACTTGTAATAAATGTA
 TATGTTTTATAATTCTGGAGAGACATAAGGAGTCTACCCGTTGAGGTTGGAGAGGGAAAAATAAGAAGCTGCCA
 CTAACAGGAGTCAACCAGGAAAGCACCGCACAGGCTGGCGCGGGACAGACTCCTAACCTGGGGCTCTGCAGTG
 GCAGGCGAGGCTGCAGGAGGCCACAGATAAGCTGGCAAGAGGAAGGATCCAGGCACATGGTTCATCACGAGCA
 TGAGGGAACAGCAAGGGGCACGGTATCACAGCCTGGAGACACCCACAGATGGCTGGATCCGGTGTACGGGAA
 ACATTTTCTAAGATGCCATGTAGAACAGACCAAGATGTGTACAGCACTATGAGCATTAATAAACCTTCCAGAAT
 CAATAATCCGTGGCAACATATCTGTAAAAACAAACACTGTAACCTCTAAATAAATGTTTAGTCTTCCCTGTAAAA

FIGURE 36

MLRGTMTAWRGMRPEVTLACLLLATAGCFADLNEVPQVTVQPASTVQKPGGTVILGCVVEPP
 RMNVTWRLNGKELNGSDDALGVLITHGTLVITALNNHTVGRYQCVARMPAGAVASVPATVTL
 ANLQDFKLDVQHVRVIEVDEGNTAVIACHLPESHHPKAQVRYSVKQEWLEASRGNYLIMPSGNLQ
 IVNASQEDEGMYKCAAYNPVTQEVKTSGSSDRLRVRRSTAEAAARI IYPPEAQTIIVTKGQSL
 ILECVASGI PPPRVTWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGA
 AVILYNVQVFEPPEVTMELS QLVIPWGQSAKLTCEVRGNPPPSVLWLRNAVPLISSQRLRLS
 RRALRVLSMGPEDEGVYQCMAENEVGSAAHAVVQLRTSRPSITPRLWQDAELATGTPPVSPSK
 LGNPEQMLRGQPALPRPPTSVGPASPKCPGEGKGQGAPEAPIILSSPRTSKTDSYELVWRPR
 HEGSGRAPILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLTRLDPGSLYEVEMAAYNCAGE
 GQTAMVTFTTGRRPKPEIMASKEQQIQRDDPGAS PQSSSQPDHGRLSPPEAPDRPTISTASE
 TSVYVTWIPRNGGGFPIQSFVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGT SYKFRV
 RALNMLGESEPSAPSRPYVVS GYSGRVYERPVAGPYITFTDAVNETTIMLKWMIIPASNNNT
 PIHGFYIYYRPTDSDNDSYKKDMVEGDKYWHSISHLQPETSYDIKMQCFNEGGESEFSNVM
 ICETKARKSSGQPGRLPPPTLAPPQPLPETIERPVGTGAMVARSSDLPYLIVGVVLGSIVL
 IIVTFIPFCLWRAWKQKHTTDLGFPRSALPPSCPYTMVPLGGLPGHQASGQPYLSGISGRA
 CANGIHMNRGCPSAAVGYPGMKPQQHCPGELQQQSDTSSLLRQTHLGNGYDPQSHQITRGPK
 SSPDEGSFLYTL PDDSTHQLLQPHHDCCQRQEQPAAVGQSGVRRAPDSPVLEAVWDPPFHSG
 PPCCGLGLVPVEEVDSPDSCQVSGGDWCPQHPVGAYVGQEPGMQLSPGPLVRVSFETPPLTI

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 16-30 (type II), 854-879

FIGURE 37

CGGGAGGCTGGGTCTGTCATGATCCGGACCCCATTTGTCGGCCTCTGCCCATCGCCTGCTCCTC
CCAGGCTCCCGCGGCCGACCCCCGCGCAAC**ATG**CAGCCCACGGGCCGCGAGGGTTCCCGCGC
GCTCAGCCGGCGGTATCTGCGGCGTCTGCTGCTCCTGCTACTGCTGCTGCTGCTGCGGCAGC
CCGTAACCCGCGCGGAGACCACGCCGGGCGCCCCCAGAGCCCTCTCCACGCTGGGCTCCCCC
AGCCTCTTCACCACGCCGGGTGTCCCCAGCGCCCTCACTACCCCAGGCCTCACTACGCCAGG
CACCCCCAAAACCCTGGACCTTCGGGGTCGCGCGCAGGCCCTGATGCGGAGTTTCCCCTCG
TGGACGGCCACAATGACCTGCCCCAGGTCTGAGACAGCGTTACAAGAATGTGCTTCAGGAT
GTTAACCTGCGAAATTTAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGT
GGGTGCCCAGTTCTGGTCAGCCTCCGTCTCATGCCAGTCCCAGGACCAGACTGCCGTGCGCC
TCGCCCTGGAGCAGATTGACCTCATTACCGCATGTGTGCCTCCTACTCTGAACTCGAGCTT
GTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAGCTGGCCTGCCTCATTGGCGTGNAGGG
TGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCT
ACCTGACACTTACCTTCACCTGCAGTACACCATGGGCAGAGAGTTCCACCAAGTTCAGACAC
CACATGTACACCAACGTCAGCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAA
CCGCCTGGGCATGATGATAGATTTGTCTATGCATCGGACACCTTGATAAGAAGGGTCTGG
AAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGCTGCCAGAGCTGTGTGTGACAATTTG
TTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGACACT
GTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTG
ACCACATCAGGGCAGTCATTGGATCTGAGTTCATCGGGATTGGTGGAATTTATGACGGGACT
GGCCGGTTCCCTCAGGGGCTGGAGGATGTGTCCACATACCCAGTCCTGATAGAGGAGTTGCT
GAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTGTCTTCGTGGAAACCTGCTGCGGGTCT
TCAGACAAGTGGAAGAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCGTGGAGGCTGAGTTT
CCATATGGGCAACTGAGCACATCCTGCCACTCCCACCTCGTGCCCTCAGAATGGACACCAGGC
TACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCCTGGAGGTCCTCAAATGCCT
CCCCATACCTTGTTCCAGGCCTTGTGGCTGCTGCCACCATCCCAACCTTCACCCAGTGGCTC
TGCT**TGA**CACAGTCGGTCCCCGCGAGGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCT
AGTTCATTACAAAGCATATGCTGAGAATAAACATGTTACACATGGAAAA

FIGURE 38

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817
```

><subunit 1 of 1, 487 aa, 1 stop, 2 unknown

><MW: 53569.32, pI: 7.68, NX(S/T): 5

MQPTGREGSRALSRRYLRRLLLLLLLLLLLLLRQPVTRAETTPGAPRALSTLGSPSLFTTPGVPS
ALTTPGLTTPGTPKTLDLRGAQALMRSFPLVDGHNDLPQVLRQRYKNVLQDVNLRNFESHGQ
TSLDRLRDGLVGAQFWSASVSCQSQDQTAVRLALEQIDLIHRMCASYSELELVTSAEGLNSS
QKLACLIGVXGGHSLDSSLSVLRSFYVLGVRYLTLTFTCSTPWAESSTKFRHHMYTNVSGLT
SFGEKVVEELNRLGMMIDLSYASDTLIRRVLEVSQAPVIFSHSAARAVCDNLLNVPDDILQL
LKNGGIVMVTLSMGVLQCNLLANVSTVADHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLEDV
STYPVLIIEELLSRXWSEEELQGVLRGNLLRVFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
SHLVPQNGHQATHLEVTKQPTNRVPWRSSNASPYLVPGLVAAATIPTFTQWLC

Important features of the protein:

Signal peptide:

amino acids 1-36

Transmembrane domain:

amino acids 313-331

N-glycosylation sites.

amino acids 119-122, 184-187, 243-246 and 333-336

N-myristoylation sites.

amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359, 357-362, 394-399, 427-432 and 472-477.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 136-146

FIGURE 39

TGCTAGGCTCTGTCCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACACCCACAGA
 TCCCTCTATGACTGCAATGTGAGGTGTCCGGCTTTGCTGGCCCAGCAAGCCTGATAAGC**ATG**
 AAGCTCTTATCTTTGGTGGCTGTGGTTCGGGTGTTTGCTGGTGCCCCAGCTGAAGCCAACAA
 GAGTTCTGAAGATATCCGGTGCAAATGCATCTGTCCACCTTATAGAAACATCAGTGGGCACA
 TTTACAACCAGAATGTATCCCAGAAGGACTGCAACTGCCTGCACGTGGTGGAGCCCATGCCA
 GTGCCTGGCCATGACGTGGAGGCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAG
 CACCACCACCATCAAGGTCATCATTGTCATCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCT
 ACATGGCCTTCCTGATGCTGGTGGACCTCTGATCCGAAAGCCGGATGCATACACTGAGCAA
 CTGCACAATGAGGAGGAGAATGAGGATGCTCGCTCTATGGCAGCAGCTGCTGCATCCCTCGG
 GGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGCCCAGCAGCGGTGGAAGCTGC
 AGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGC**TAG**ATGGGCTGG
 TGTGGTTGGGTCAAGGCCCAACACCATGGCTGCCAGCTTCCAGGCTGGACAAAGCAGGGGG
 CTACTTCTCCCTTCCCTCGGTTCCAGTCTTCCCTTTAAAGCCTGTGGCATTTTTCTCTCTT
 CTCCCTAACTTTAGAAATGTTGTACTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCT
 GATCTCTGTTGTCTTCTTGGGTCTTTGGGGTTGAAGGGAGGGGGAAGGCAGGCCAGAAGGGA
 ATGGAGACATTCGAGGCGGCCTCAGGAGTGGATGCGATCTGTCTCTCCTGGCTCCACTCTTG
 CCGCCTTCCAGCTCTGAGTCTTGGGAATGTTGTTACCCTTGGAAGATAAAGCTGGGTCTTCA
 GGAATCAGTGTCTGGGAGGAAAGCATGGCCCAGCATTAGCATGTGTTCCCTTTCTGCAGTG
 GTTCTTATCACACCTCCCTCCCAGCCCCGGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTG
 AGGACAGCTCTGATGGGAGAGCTGGGCCCCCTGAGCCCACTGGGTCTTCAGGGTGCAGTGG
 AGCTGGTGTTCGCTGTCCCTGTGCACTTCTCGCACTGGGGCATGGAGTGCCCATGCATACT
 CTGCTGCCGGTCCCCCTCACCTGCACTTGAGGGGTCTGGGCAGTCCCTCCTCTCCCCAGTGTC
 CACAGTCACTGAGCCAGACGGTTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGAA
 CACCACAGCCCCTGTACTTGGGTGCGCTCTTGTCCCTGAACTTCGTTGTACCAGTGCATGGA
 GAGAAAATTTTGTCTCTTGTCTTAGAGTTGTGTGTAAATCAAGGAAGCCATCATTAATTTG
 TTTTATTCTCTCA

FIGURE 40

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278

<subunit 1 of 1, 183 aa, 1 stop

<MW: 20574, pI: 6.60, NX(S/T): 3

MKLLSLVAVVGCLLVPPAEANKSSEDIRCKCICPPYRNISGHIYNQNVSQKDCNCLHVVEPM
PVPBGHDVEAYCLLCECRYEERSTTTIKVIIIVYLSVVGALLLYMAFLMLVDPLIRKPDAYTE
QLHNEEENEDARSMAAAAASLGGPRANTVLERVEGAQQRWKLQVQEQRKTVFDRHKMLS

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 90-112

N-glycosylation sites.

amino acids 21-24, 38-41 and 47-50

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 41

AGCGGGTCTCGCTTGGGTTCCGCTAATTTCTGTCTGAGGCGTGAGACTGAGTTCATAGGGTCTGGGTCCCCGA
 ACCAGGAAGGGTTGAGGGAACACAATCTGCAAGCCCCCGGACCCCAAGTGAGGGGCCCCGTGTTGGGGTCTCCC
 TCCCTTTGCATTCCCACCCCTCCGGGCTTTGCGTCTTCTGGGGACCCCTCGCCGGGAGATGGCCGCGTTGATG
 CGGAGCAAGGATTCGTCTGCTGCCTGCTCCTACTGGCCGCGGTGCTGATGGTGGAGAGCTCACAGATCGGCAGT
 TCGCGGGCCAACTCAACTCCATCAAGTCTCTCTGGGCGGGGAGACGCCTGGTCAGGCCGCCAATCGATCTGCG
 GGCATGTACCAAGGACTGGCATTGCGCGGCAGTAAGAAGGGCAAAAACCTGGGGCAGGCCTACCCCTGTAGCAGT
 GATAAGGAGTGTGAAGTTGGGAGGTATTGCCACAGTCCCCACCAAGGATCATCGGCCTGCATGGTGTGTCGGAGA
 AAAAAGAAGCGCTGCCACCGAGATGGCATGTGCTGCCCCAGTACCCGCTGCAATAATGGCATCTGTATCCCAGTT
 ACTGAAAGCATCTTAACCCCTCACATCCCGCTCTGGATGGTACTCGGCACAGAGATCGAAACCACGGTCATTAC
 TCAAACCATGACTTGGGATGGCAGAATCTAGGAAGACCACACACTAAGATGTCACATATAAAAGGGCATGAAGGA
 GACCCCTGCCTACGATCATCAGACTGCATTGAAGGGTTTTGCTGTGCTCGTCATTTCTGGACCAAATCTGCAAA
 CCAGTGTCTCCATCAGGGGGAAGTCTGTACCAAACAACGCAAGAAGGGTTCTCATGGGCTGGAAATTTTCCAGCGT
 TCGCACTGTGCGAAGGGCCTGTCTTGCAAAGTATGGAAGATGCCACCTACTCCTCCAAAGCCAGACTCCATGTG
 TGTCAGAAAAATTTGATCACCATTGAGGAACATCATCAATTGCAGACTGTGAAGTTGTGTATTTAATGCATTATAG
 CATGGTGGAAAAATAAGGTTTCAAGATGCAGAAGAATGGCTAAAATAAGAAACGTGATAAGAATATAGATGATCACA
 AAAGGGAGAAAGAAAACATGAAGTGAATAGATTAGAATGGGTGACAAATGCAGTGCAGCCAGTGTTCCTATTATG
 CAACTTGTCTATGTAAATAATGTACACATTTGTGGAAAATGCTATTATTAAGAGAACAAGCACACAGTGGAAAT
 ACTGATGAGTAGCATGTGACTTTCCAAGAGTTTAGGTTGTGCTGGAGGAGAGGTTTCTTCCAGATTGCTGATTGC
 TTATACAAATAACCTACATGCCAGATTTCTATTCAACGTTAGAGTTTAAACAAAATACTCCTAGAATAACTTGTTA
 TACAATAGGTTTCTAAAATAAAAATTGCTAAACAAGAAATGAAACATGGAGCATTGTTAATTTACAACAGAAAAAT
 TACCTTTTGTATTTGTAACACTACTTCTGCTGTTCAATCAAGAGTCTTGGTAGATAAGAAAAAATCAGTACATAT
 TTCCAAATAAATTGCAAATAATGGCCAGTTGTTTAGGAAGGCCTTTAGGAAGACAAATAAATAACAAACAAACAG
 CCACAAATACTTTTTTTTTTCAAATTTTAGTTTTACCTGTAATTAAGAAGTATACAAAGACAAAAACAGTTCC
 TTCAGATTCTACGGAATGACAGTATATCTCTTTATCCTATGTGATTCCTGCTCTGAATGCATTATATTTTCCA
 AACTATACCCATAAATTGTGACTAGTAAATACTTACACAGAGCAGAATTTTACAGATGGCAAAAAAATTTAAA
 GATGTCCAATATATGTGGGAAAAGAGCTAACAGAGAGATCATTATTTCTTAAAGATTGGCCATAACCTATATTTT
 GATAGAATTAGATTGGTAAATACATGTATTCTACATACTCTGTGGTAATAGAGACTTAAGCTGGATCTGTACTG
 CACTGGAGTAAGCAAGAAAATTGGGAAAACCTTTTTCGTTTGTTCAGGTTTTGGCAACACATAGATCATATGTCTG
 AGGCACAAGTTGGCTGTTTCATCTTTGAAACCAGGGGATGCACAGTCTAAATGAATATCTGCATGGGATTTGCTAT
 CATAATATTTACTATGCAGATGAATTCAGTGTGAGGTCTGTGTCCTACTATCCTCAAATTTATTTTATAG
 TGCTGAGATCCTCAAATAATCTCAATTTAGGAGGTTTCAAAAAATGACTCCTGAAGTAGACAGAGTAGTGAGG
 TTTTATTGCCCTCTATAAGCTTCTGACTAGCCAATGGCATCATCCAATTTTCTTCCCAACCTCTGCAGCATCTG
 CTTTATTGCCAAAGGGCTAGTTTCCGTTTCTGCGAGCATTGCGGTTAAAAAATATAAGTAGGATAAATTGTAAA
 ACCTGCATATTGCTAATCTATAGACACCACAGTTTCTAAATTTCTTGAACCACTTTACTACTTTTTTTAACTT
 AACTCAGTTCTAAATACTTTGTCTGGAGCACAAAAAATAAAGGTTATCTTATAGTCGTGACTTTAACTTTTG
 TAGACCACAATTCACTTTTTAGTTTTCTTTTACTTAAATCCCATCTGCAGTCTCAAATTTAAGTTCTCCAGTAG
 AGATTGAGTTTGAGCCTGTATATCTATTAAAAATTTCAACTTCCCACATATATTTACTAAGATGATTAAGACTTA
 CATTTTCTGCACAGGTCTGCAAAAAAATAAATAAAGGTTATCTTATAGTCGTGACTTTAACTTTTG
 TGCTATAAGCTTGTGAAATGAAATGGAACATTTCAATCAACATTTCTATATAACAATTATTATATTTACAAAT
 TTGGTTTCTGCAATATTTTCTTATGTCCACCCTTTTAAAAATTATTATTTGAAGTAATTTATTTACAGGAAATG
 TTAATGAGATGTATTTTCTTATAGAGATATTTCTTACAGAAAGCTTTGTAGCAGAATATATTTGCAGCTATTGAC
 TTTGTAATTTAGGAAAAATGTATAAATAAGATAAAATCTATTAAATTTTCTCCTCTAAAAACTGAAAAAATAA
 AAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

MAALMRSKDSSCCLLLLAAVLMVESSQIGSSRAKLNSIKSSLGGETPGQAANRSAGMYQGLA
FGGSKKGKNLGQAYPCSSDKECEVGRYCHSPHQSSACMVCRRKKKRCHRDGMCCPSTRCNN
GICIPVTESILTPHIPALDGTRHRDRNHGHYSNHDLGWQNLGRPHTKMSHIKGHEGDPCLRS
SDCIEGFCCARHFWTKICKPVLHQGEVCTKQRKKGSHGLEIFQRCDCAKGLSCKVWKDATYS
SKARLHVCQKI

Signal peptide:

amino acids 1-25

Parameter	Unit	Value	Standard Error	t-Statistic	p-Value
Intercept		1.0000	0.0000	1.0000	0.0000
Age	Years	0.0000	0.0000	0.0000	0.0000
Gender		0.0000	0.0000	0.0000	0.0000
Marital Status		0.0000	0.0000	0.0000	0.0000
Education	Years	0.0000	0.0000	0.0000	0.0000
Income	Thousands of Dollars	0.0000	0.0000	0.0000	0.0000
Health		0.0000	0.0000	0.0000	0.0000
Smoking		0.0000	0.0000	0.0000	0.0000
Alcohol		0.0000	0.0000	0.0000	0.0000
Exercise		0.0000	0.0000	0.0000	0.0000
Stress		0.0000	0.0000	0.0000	0.0000
Family Size		0.0000	0.0000	0.0000	0.0000
Home Ownership		0.0000	0.0000	0.0000	0.0000
Commuting Time	Hours	0.0000	0.0000	0.0000	0.0000
Job Satisfaction		0.0000	0.0000	0.0000	0.0000
Work-Life Balance		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Overall Health		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.0000	0.0000	0.0000	0.0000
Life Satisfaction		0.0000	0.0000	0.0000	0.0000
Life Expectancy	Years	0.0000	0.0000	0.0000	0.0000
Quality of Life		0.			

FIGURE 43

TCTCAATCTGCTGACCTCGTGATCCGCCTGACCTTGTAAATCCACCTACCTTGGCCTCCCAA
 GTGTTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAACATCACGTTTTTAAAAATTGATT
 TCTTCAAATTCATGGCAAATATTTCCCTTCCCTTTAACTTCTTATGTCAGAATGAGGAAGGA
 TAGCTGCATTTATTTAGTCAGTTTTTCATTGCATAGTAATATTTTCATGTAGTATTTTCTAAG
 TTATATTTTAGTAATTCATATGTTTTAGATTATAGGTTTTAACATACTTGTGAAAATACTTG
ATGTGTTTTTAAAGCCTTGGGCAGAAATTCTGTATTGTTGAGGATTTGTTCTTTTTATCCCCCT
 TTTAAAGTCATCCGTCCTTGGCTCAGGATTTGGAGAGCTTGCACCACCAAAAATGGCAAACA
 TCACCAGCTCCCAGATTTTGGACCAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACC
 CCAAGTACACAGCAGAATAGTACAAGTCACCCTACAATACTACTTCTTGGGACCTCAAGCC
 CCCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAAATCTCAACCTGAGCCATCCC
 CAGTTCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCACTGTTCTCTCT
 CCTGGTTTGGAGTCCTTTCTTCCCAGGCAAACTTCGAGAATCAACACCTGGAGACAGTCC
 CTCCACTGTGAACAAGCTTTTGCAGCTTCCCAGCACGACCATTGAAAATATCTCTGTGTCTG
 TCCACCAGCCACAGCCCAAACACATCAAACCTTGCTAAGCGGCGGATACCCCCAGCTTCTAAG
 ATCCCAGCTTCTGCAGTGGAAATGCCTGGTTTCAGCAGATGTCACAGGATTAAATGTGCAGTT
 TGGGGCTCTGGAATTTGGGTCAGAACCTTCTCTCTCTGAATTTGGATCAGCTCCAAGCAGTG
 AAAATAGTAATCAGATTCCCATCAGCTTGTATTCTGAAGTCTTTAAGTGAGCCTTTGAATACA
 TCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAACCTTCCGTCATTACCTCCTG
 CAGTCTGACAAGCTCATCACTGAATTCTGCTAGTCCAGTAGCAATGTCTTCCTCTTATGACC
 AGAGTTCTGTGCATAACAGGATCCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGTCCA
 GGAACCATCATGAATGGACATGGTGGTGGTCGAAGTCAGCAGACACTAGACAGTAAGTATAG
 CAGCAAGCTACTCTTGTTCATGGCTGGTGCCAACCAAACAGAGGAAGAGGATAGCTCACGTGA
 TGTGGAAAACACCAGTTGGTCAATGGCTCATTCGT**TAAAA**AGCAGCCCTTTTGCTTTTTTGT
 TTTTGGACCAGGTGTTGGCTGTGGTGTATTAGAAATGTCTTAACCACAGCAAGAAGGAGGT
 GGTGGTCTCATATTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATAACAGTATGCAT
 TTTAAAGATGCTTGGGCCAGGCGGGGTGGCTGATGCCCATAAATCCCAGTGCTTTGGGGGGCC
 AAGGCAGGCAGATTGCCCAAGCTCAGGAGTTTGGAGACCACCCTGGGCAACATGGTGAAACTC
 TGTCTCTACTAAAATACGAAAACTAGCCGGGTGTGGTGGCGGCGCGTGCCTGTAATCCCAG
 CTAATTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCC
 GTCTGAAAAGA

T
C
T
C
A
A
T
C
T
G
C
T
G
A
C
C
T
C
G
T
G
A
T
C
C
G
C
C
T
G
A
C
C
T
T
G
T
A
A
T
C
C
A
C
C
T
A
C
C
T
T
G
G
C
C
T
C
C
C
A
A

FIGURE 44

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTT
PSTQQNSTSHPTTTTSDWDLKPPTSQSSVLSHLDFKSQPEPSPVLSQLSQRQQHQSQAQTVPP
PGLESFPSQAKLRESTPGDSPSTVNKLLQLPSTTIENISVSVHQPPKHIKLAKRRIPPASK
IPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAPSSSENSNQIPISLYSKSLSEPLNT
SLSMTSAVQNSTYTTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPIYQSPVSSSESAP
GTIMNGHGGGRSQQTLDISKYSSKLLLSWLVPKQKRIAHVMWKTPVGQWLIR

Signal peptide:

amino acids 1-24

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 45

GCCGAGTGGGACAAAGCCTGGGGCTGGGCGGGGGCC**ATG**GCGCTGCCATCCCGAATCCTGCT
 TTGGAAACTTGTGCTTCTGCAGAGCTCTGCTGTTCTCCTGCACTCAGCGGTGGAGGAGACGG
 ACGCGGGGCTGTACACCTGCAACCTGCACCATCACTACTGCCACCTCTACGAGAGCCTGGCC
 GTCCGCCTGGAGGTCACCGACGGCCCCCGGCCACCCCGCCTACTGGGACGGCGAGAAGGA
 GGTGCTGGCGGTGGCGCGGGCGCACCCGCGCTTCTGACCTGCGTGAACCGCGGGGCACGTGT
 GGACCGACCGGCACGTGGAGGAGGCTCAACAGGTGGTGCCTGGGACCGGCAGCCGCCCGGG
 GTCCCGCACGACCGCGCGGACCGCCTGCTGGACCTCTACGCGTCGGGCGAGCGCCGCGCCTA
 CGGGCCCCCTTTTTCTGCGCGACCGCGTGGCTGTGGGCGCGGATGCCTTTGAGCGCGGTGACT
 TCTCACTGCGTATCGAGCCGCTGGAGGTCGCCGACGAGGGCACCTACTCCTGCCACCTGCAC
 CACCATTACTGTGGCCTGCACGAACGCCGCGTCTTCCACCTGACGGTCGCCGAACCCACGC
 GGAGCCGCCCCCGGGGCTCTCCGGGCAACGGCTCCAGCCACAGCGGCGCCCCAGGCCCAG
 ACCCCACACTGGCGCGCGGCCACAACGTCATCAATGTCATCGTCCCCGAGAGCCGAGCCCAC
 TTCTTCCAGCAGCTGGGCTACGTGCTGGCCACGCTGCTGCTCTTCATCCTGCTACTGGTCAC
 TGTCTCCTGGCCGCCCGCAGGCGCCGCGGAGGCTACGAATACTCGGACCAGAAGTCGGGAA
 AGTCAAAGGGGAAGGATGTAACTTGGCGGAGTTCGCTGTGGCTGCAGGGGACCAGATGCTT
 TACAGGAGTGAGGACATCCAGCTAGATTACAAAAACAACATCCTGAAGGAGAGGGCGGAGCT
 GGCCACAGCCCCCTGCCTGCCAAGTACATCGACCTAGACAAAGGGTTCCGGAAGGAGAACT
 GCAAA**TAG**GGAGGCCCTGGGCTCCTGGCTGGGCCAGCAGCTGCACCTCTCCTGTCTGTGCTC
 CTCGGGGCATCTCCTGATGCTCCGGGGCTCACCCCCCTTCCAGCGGCTGGTCCCGCTTTCCT
 GGAATTTGGCCTGGGCGTATGCAGAGGCCGCTCCACACCCCTCCCCAGGGGCTTGGTGGC
 AGCATAGCCCCACCCCTGCGGCCTTTGCTCACGGGTGGCCCTGCCACCCCTGGCACAACC
 AAAATCCCCTGATGCCATCATGCCCTCAGACCCTTCTGGGCTCTGCCCCTGGGGGCTG
 AAGACATTCTGGAGGACACTCCCATCAGAACCTGGCAGCCCCAAAACCTGGGGTCAGCCTCA
 GGGCAGGAGTCCCACTCCTCCAGGGCTCTGCTCGTCCGGGGCTGGGAGATGTTCTGGAGGA
 GGACACTCCCATCAGAACTTGGCAGCCTTGAAGTTGGGGTCAGCCTCGGCAGGAGTCCCACT
 CCTCCTGGGGTGCTGCCTGCCACCAAGAGCTCCCCACCTGTACCACCATGTGGGACTCCAG
 GCACCATCTGTTCTCCCCAGGGACCTGCTGACTTGAATGCCAGCCCTTGCTCCTCTGTGTTG
 CTTTGGGCCACCTGGGGCTGCACCCCTGCCCTTTCTCTGCCCCATCCCTACCCTAGCCTTG
 CTCTCAGCCACCTTGATAGTCACTGGGCTCCCTGTGACTTCTGACCCTGACACCCCTCCCTT
 GGACTCTGCCTGGGCTGGAGTCTAGGGCTGGGGCTACATTTGGCTTCTGTACTGGCTGAGGA
 CAGGGGAGGGAGTGAAGTTGGTTTGGGGTGGCCTGTGTTGCCACTCTCAGCACCCACATTT
 GCATCTGCTGGTGGACCTGCCACCATCACAATAAAGTCCCCATCTGATTTTAAAAAAAAA
 AAAAA

FIGURE 46

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60618

<subunit 1 of 1, 341 aa, 1 stop

<MW: 38070, pI: 6.88, NX(S/T): 1

MALPSRILLWKLVLQSSAVLLHSAVEETDAGLYTCNLHHHYCHLYESLAVRLEVTDGPPAT
PAYWDGEKEVLAVARGAPALLTCVNRGHVWTDHRVVEEAQQVVHWRQPPGVPHDRADRLDL
YASGERRAYGPLFLRDRVAVGADAFERGFSLRIEPLVADEGTYSCHLHHHYCGLHERRVF
HLTVAEPHAEPPIRGSPGNGSSHSGAPGPDPTLARGHNVINVIVPESRAHFFQQLGYVLATL
LLFILLLVTVLLAARRRRGGYEYSDQKSGKSKGKDVNLAEFAVAAGDQMLYRSEDIQLDYKN
NILKERAELAHSPLPAKYIDLDKGFRKENCK

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 237-262

N-glycosylation site.

amino acids 205-208

Cell attachment sequence.

amino acids 151-154

Coproporphyrinogen III oxidase proteins.

amino acids 115-140

FIGURE 47

CGCCGGAGGCAGCGGGCGGCGTGGCGCAGCGGGCAGC**ATG**GCCGTTGTCTCAGAGGACGACTTT
CAGCACAGTTCAAACCTCCACCTACGGAACCACAAGCAGCAGTCTCCGAGCTGACCAGGAGGC
ACTGCTTGAGAAGCTGCTGGACCGCCCGCCCCCTGGCCTGCAGAGGCCCCGAGGACCGCTTCT
GTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCAGTCTACTGCCATGGAACCTTCTTT
ATCACTGCCAAGGAGTACTGGATGTTCAAACCTCCGCAACTCCTCCAGCCCAGCCACCGGGGA
GGACCCTGAGGGCTCAGACATCCTGAACTACTTTGAGAGCTACCTTGCCGTTGCCTCCACCG
TGCCCTCCATGCTGTGCCTGGTGGCCAACCTTCCTGCTTGTCAACAGGGTTGCAGTCCACATC
CGTGTCTTGGCCTCACTGACGGTCATCCTGGCCATCTTCATGGTGATAACTGCACTGGTGAA
GGTGGACACTTCCTCCTGGACCCGTGGTTTTTTTTGCGGTCACCATTTGTCTGCATGGTGATCC
TCAGCGGTGCCTCCACTGTCTTCAGCAGCAGCATCTACGGCATGACCGGCTCCTTTCTCTATG
AGGAACTCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCGTGGCCTC
ATTGGTGGACTTGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCCTGACGG
CCACCATCTTCCTCGTGCTCTGCATGGGACTCTACCTGCTGCTGTCCAGGCTGGAGTATGCC
AGGTACTACATGAGGCCTGTTCTTGCGGCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCCA
GGACTCCCTCAGTGCCCCCTTCGGTGGCCTCCAGATTCAATTGATTCCACACACCCCCCTCTCC
GCCCCATCCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTACCTACGTCTTCTTCATCACC
AGCCTCATCTACCCCGCCGTCTGCACCAACATCGAGTCCCTCAACAAGGGCTCGGGCTCACT
GTGGACCACCAAGTTTTTCATCCCCCTCACTACCTTCCTCCTGTACAACCTTTGCTGACCTAT
GTGGCCGGCAGCTCACCGCCTGGATCCAGGTGCCAGGGCCCAACAGCAAGGCGCTCCCAGGG
TTCGTGCTCCTCCGGACCTGCCTCATCCCCCTCTTCGTGCTCTGTAACCTACCAGCCCCGCGT
CCACCTGAAGACTGTGGTCTTCCAGTCCGATGTGTACCCCGCACTCCTCAGCTCCCTGCTGG
GGCTCAGCAACGGCTACCTCAGCACCTGGCCCTCCTCTACGGGCCTAAGATTGTGCCCAGG
GAGCTGGCTGAGGCCACGGGAGTGGTGATGTCCTTTTATGTGTGCTTGGGCTTAACACTGGG
CTCAGCCTGCTCTACCTCCTGGTGCACCTCATC**TAGA**AAGGGAGGACACAAGGACATTGGTG
CTTCAGAGCCTTTGAAGATGAGAAGAGAGTGCAGGAGGGCTGGGGGCCATGGAGGAAAGGCC
TAAAGTTTTCACTTGGGGACAGAGAGCAGAGCACACTCGGGCCTCATCCCTCCCAAGATGCCA
GTGAGCCACGTCCATGCCATTCCGTGCAAGGCAGATATTCCAGTCATATTAACAGAACACT
CCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACAGCTGATGGTTA
ACATTCCACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTCGCCCTAGAGTTATTACA
AAGCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCAGCTGCGCTCATTCCAGCT
GACAGCGAGATGCAAGCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGA
AGTCCCCTGGCATGGTCAGTCTCAGGCCCAAGACTCAAGTGTGCACAGACCCCTGTGTTCT
GCGGGTGAACAACTGCCCACTAACCAGACTGGAAAACCCAGAAAGATGGGCCTTCCATGAAT
GCTTCATTCCAGAGGGACCAGAGGGCCTCCCTGTGCAAGGGATCAAGCATGTCTGGCCTGGG
TTTTCAAAAAAAGAGGGATCCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTC
TTTCAGTGTTCTGTTTACAACATGTCAAAGCCATTGGTTCAAGGGCGTAATAAATACTTGC
GTATTCAAAAA

FIGURE 48

MAVSEDDDFQHSSNSTYGTSSSLRADQEALLEKLLDRPPPGLRPEDRFCGTYYIIFSLGI
GSLLPWNFFITAKEYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFL
LVNRVAVHIRVLASLTVILAI FMVITALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSI
YGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSDVRNSALAFFLTATIFLVLCMGly
LLLSRLEYARYYMRPVLAHVFSGEEELPQDSL SAPSVASRFIDSHTPPLRPILKKTASLGF
CVTYVFFITS LIYPAVCTNIESLNKGSGLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVP
GPNSKALPGFVLLRTCLIPLFVLCNYQPRVHLKTVVVFQSDVYPALLSSLLGLSNGYLSTLAL
LYGPKIVPRELAEATGVVMSFYVCLGLTLGSACSTLLVHLI

Transmembrane domain:

amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252,
305-330, 448-472

protein sequence

FIGURE 50

MAPLALHLLVLPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVVWGLNRTLKPQRV
IVVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRII
HKLCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGALRPQEKGHSPEDIYQMA
LNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSEDGEFFYLSFAEALR
AHSCLSDRLQYSRIVGGWDLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNL
KVLKADVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGG
HSNTDRPSRMIFYPPPREGALLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQL
WDGTGVVKRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKS
ALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQ
NTTHTRTSH

Signal peptide:

amino acids 1-21

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 51

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTGCCTGCAT
GGACGCTCTGAAGCCACCCTGTCTCTGGAGGAACCACGAGCGAGGGAAGGACAGGGACTCGTGTGGCAGGAA
GAACTCAGAGCCGGGAAGCCCCATTCACTAGAAGCACTGAGAGATGCGGCCCTTCGCAGGGTCTGAATTTCT
GCTGCTGTTTACAAAGATGCTTTTTATCTTTAACTTTTTGTTTTCCCACTTCCGACCCCGGCGTTGATCTGCAT
CCTGACATTTGGAGCTGCCATCTTCTTGTGGCTGATCACCAGACCTCAACCCGTCTTACCTCTTCTTGACCTGAA
CAATCAGTCTGTGGGAATTGAGGGAGGAGCACGGAAGGGGGTTTCCAGAGAACAATGACCTAACAAGTTGCTG
CTTCTCAGATGCCAAGACTATGTATGAGTTTTCCAAAGAGGACTCGCTGTGTCTGACAATGGGCCCTGCTTGGG
ATATAGAAAACCAACCAGCCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTT
CTGTCTCTTGATATAAAGTTATAAATCATCACCAGACAGTTTGTGCGCATCTTTGCTCAGAATAGGCCAGAGTG
GATCATCTCCGAATTGGCTTGTTACACGTACTCTATGGTAGCTGTACCTCTGTATGACACCTTGGGACCAGAAGC
CATCGTACATATTGTCAACAAGGCTGATATCGCCATGGTGATCTGTGACACACCCCAAGGCATTGGTGCTGAT
AGGGAATGTAGAGAAAGGCTTACCCCGAGCCTGAAGGTGATCATCCTTATGGACCCCTTTGATGATGACCTGAA
GCAAAGAGGGGAGAAGAGTGGAATTGAGATCTTATCCCTATATGATGCTGAGAACCTAGGCAAAGAGCACTTCA
AAAACCTGTGCCTCCTAGCCCAGAAGACCTGAGCGTCATCTGCTTACCAGTGGGACCACAGGTGACCCCAAAGG
AGCCATGATAACCCATCAAAATATTGTTTTCAAATGCTGCTGCCCTTCTCAAATGTGTGGAGCATGCTTATGAGCC
CACTCCTGATGATGTGGCCATATCCTACCTCCCTCTGGCTCATATGTTTGAGAGGATTGTACAGGCTGTTGTGTA
CAGCTGTGGAGCCAGAGTTGGATTCTTCCAAGGGGATATTCGGTTGCTGGCTGACGACATGAAGACTTTGAAGCC
CACATTGTTTTCCCGCGGTGCCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCCTT
GAAGAAGTTCTTGTTGAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAGGGTATCATCAGGCATGATAG
TTTTCTGGGACAAGCTCATCTTTGCAAAGATCCAGGACAGCCTGGGCGGAAGGTTCTGTAAATTGCTACTGGAGC
TGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTGAGGTGTATGAAGCTTATGGTCA
AACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTCACGTTGGGGTGCCCTGGC
TTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAACACTTTTACAGTGAATAATGAAGGAGAGGTCTGCAT
CAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAGAAGACACAGGAAGCCCTGGACAGTGATGGCTG
GCTTCACACAGGAGACATTGGTTCGCTGGCTCCCGAATGGAACCTCTGAAGATCATCGACCGTAAAAAGAACATTTT
CAAGCTGGCCCAAGGAGAATACATTGCACCAGAGAAGATAGAAAATATCTACAACAGGAGTCAACCAGTGTTACA
AATTTTTGTACACGGGGAGAGCTTACGGTCATCCTTAGTAGGAGTGGTGGTTCCTGACACAGATGTACTTCCCTC
ATTTGCAGCCAAGCTTGGGGTGAAGGGCTCCTTTGAGGAACGTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTT
AGAAGACTTGCAGAAAATTGGGAAAGAAAGTGGCCTTAAAACCTTTTGAACAGGTCAAAGCCATTTTTCTTCATCC
AGAGCCATTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTCCAAATACTT
TCGGACCCAAATTGACAGCCTGTATGAGCACATCCAGGATTAGGATAAGGTACTTAAGTACCTGCCGGCCCACTG
TGCACTGCTTGTGAGAAAATGGATTAAAAACTATTCTTACATTTGTTTTGCCTTTCCTCCTATTTTTTTTTTAAACC
TGTTAAACTCTAAAGCCATAGCTTTTGTATATTGAGACATATAATGTGTAACTTAGTTCCCAAATAAATCA
ATCCTGTCTTTCCCATCTTCGATGTTGCTAATATTAAGGCTTCAGGGCTACTTTTATCAACATGCCTGTCTTCAA
GATCCCAGTTTATGTTCTGTGTCCTTCCCTCATGATTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGT
CAAAGGGACCCCTCTGTGCCTTCTTCTTTGTTTTGTGATAAACATAACTTGCCAACAGTCTCTATGCTTATTTACA
TCTTCTACTGTTCAAACCTAAGAGATTTTTAAATTCTGAAAACTGCTTACAATTCTATGTTTTCTAGCCACTCCAC
AAACCACTAAAATTTAGTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTT
CTGCGTAAATTAATTTGTGTACTGAAGGGAAAAGTTTGATCATACCAACATTTTCTTAACTCTCTAGTTAGATA
TCTGACTTGGGAGTATTAATAATTGGGTCTATGACATACTGTCCAAAGGAATGCTGTTCTTAAAGCATTATTTA
CAGTAGGAACCTGGGGAGTAAATCTGTTCCCTACAGTTTGTGCTGAGCTGGAAGCTGTGGGGGAAGGAGTTGACA
GGTGGGCCCAGTGAACCTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACCTCCTGAACTGGGAACAAAGATCT
ACAGGCAAGCAAGATGCCACACAACAGGCTTATTTCTGTGAAGGAACCAACTGATCTCCCCACCCTTGGATT
AGAGTTCCTGCTCTACCTTACCCACAGATAACACATGTTGTTTCTACTTGTAAATGTAAAGTCTTTAAATAAAC
TATTACAGATAAAAAA

FIGURE 52

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60775

<subunit 1 of 1, 739 aa, 1 stop

<MW: 82263, pI: 7.55, NX(S/T): 3

MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFLLLFTKMLFIFN
 FLFSPLPTPALICILTFGAAIFLWLITRPQPVLPLLDLNNQSVGIEGGARKGVSQKNNDLTS
 CCFSDAKTMYEVFQRLAVSDNGPCLGYRKPNQPYRWLSYQVSDRAEYLGSCLLHKGYKSS
 PDQFVGIFAQNRPEWIISELACYTYSMVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKAL
 VLIGNVEKGFTPSLKVIIILMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPSPED
 LSVICFTSGTTGDPKGAMITHQNIVSNAAFLKCVEHAYEPTPDDVAISYLPLAHMFERIVQ
 AVVYSCGARVGGFFQGDIRLLADDMKTLKPTLFPAVPRLNRIYDKVQNEAKTPLKKFLLKLA
 VSSKFKEQLQKGIIRHDSFWDKLIFAKIQDSLGGRRVIVITGAAPMSTSVMTFFRAAMGCQVY
 EAYGQTECTGGCTFTLPGDWTSGHVGVPACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFK
 GYLKDPEKTQEALDSDGWLHTGDIGRWLPNGTLKIIDRKKNIFKLAQGEYIAPEKIENIYNR
 SQPVLQIFVHGESLRSSLVGVVVPD TDVLP SFAAKLG VKGSFEELCQNQVVREAILEDLQKI
 GKESGLKTFEQVKAIFLHPEPFSIENGLLTPTLKAARGELSKYFRTQIDSLYEHIQD

Important features:

Type II transmembrane domain:

amino acids 61-80

Putative AMP-binding domain signature.

amino acids 314-325

N-glycosylation site.

amino acids 102-105, 588-591 and 619-622

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 53

GGAGGCGGAGGCCGCGGCGAGCCGGGCGGAGCAGTGAGGGCCCTAGCGGGGGCCCGAGCGGGG
CCCGGGGGCCCTAAGCCATTCTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAAT
CCGGT**ATG**GACGACTGGAAGCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGG
AGCTGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGCCCTGCGGAGATTCTGTCA
GACAGGGGGCCGTGCTTTTCCTGCTGGTGACTGTCATTGTCAATATCAAGTTGATCCTGGACA
CTCGGCGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGAGCAAGACTATGATGAGGCC
CTAGGCCGCTGGAGCCCCACGGCGCAGAGGCCAGTGTTCCCGGGCGGGTCTGGACGTAGA
GGTGTATTCAAGTCGCAGCAAAGTATATGTGGCAGTGAGTGGCACCACGGTGCTGGAGGATG
AGGCCCCGGGAGCAGGGCCGGGGCATCCATGTCATTGTCTCAACCAGGCCACGGGCCACGTG
ATGGCAAAACGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGCTATTCTT
CAACATGGTAGCGCCCGGCCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACC
TCAAGGACACAGCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGC
TGGAGGGACACATGGGCCTTCGTGGGACGAAAAGGAGGTCCTGTCTTCGGGGAGAAACATTTC
TAAGTCACCTGCCCTCTCTTCTGGGGGACCCAGTCCTGCTGAAGACAGATGTGCCATTGA
GCTCAGCAGAAGAGGCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGC
TTCTGCAGCAAAGTTGAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCACACCCATCGA
GTTTCAGCCCTGACCCACTCCCAGACAACAAGGTCTCAATGTGCCTGTGGCTGTCATTGCAG
GGAACCGACCCAATTACCTGTACAGGATGCTGCGCTCTCTGCTTTCAGCCCAGGGGGTGTCT
CCTCAGATGATAACAGTTTTTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGGCACT
GTTTGGTCTGAGGGGCATCCAGCATACTCCCATCAGCATCAAGAATGCCCGCGTGTCTCAGC
ACTACAAGGCCAGCCTCACTGCCACTTTCAACCTGTTTCCGGAGGCCAAGTTTGCTGTGGTT
CTGGAAGAGGACCTGGACATTGCTGTGGATTTTTTTCAGTTTCCTGAGCCAATCCATCCACCT
ACTGGAGGAGGATGACAGCCTGTACTGCATCTCTGCCTGGAATGACCAGGGGTATGAACACA
CGGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCCTGGGCTGGGCTGGGTGCTC
AGGAGGTCTTGTACAAGGAGGAGCTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCTGGGA
TTGGGACATGTGGATGCGGATGCCTGAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACG
TTTCCCGATCCTACCACTTTGGCATCGTCGGCCTCAACATGAATGGCTACTTTCACGAGGCC
TACTTCAAGAAGCACAAGTTCAACACGGTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCT
GAAGAAAGAAGCTTATGAAGTGGAAAGTTCACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACC
ACAGCAAGAACCCTTGTGAAGACTCTTTCCTGCCAGACACAGAGGGCCACACCTACGTGGCC
TTTATTCGAATGGAGAAAGATGATGACTTCACCACCTGGACCCAGCTTGCCAAGTGCCTCCA
TATCTGGGACCTGGATGTGCGTGGCAACCATCGGGGCCTGTGGAGATTGTTTCGGAAGAAGA
ACCACTTCTGGTGGTGGGGGTCCCGGCTTCCCCCTACTCAGTGAAGAAGCCACCCTCAGTC
ACCCCAATTTTCTGGAGCCACCCCCAAAGGAGGAGGGAGCCCCAGGAGCCCCAGAACAGAC
ATGAGACCTCCTCCAGGACCCTGCGGGGCTGGGTACTGTGTACCCCCAGGCTGGCTAGCCCT
TCCCTCCATCCTGTAGGATTTTGTAGATGCTGGTAGGGGCTGGGGCTACCTTGTTTTTAACA
TGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCCCTGCTCCAACACCCCGTTCTGAGTT
AAAAGTCTATTTATTTACTTCTTGTGGAGAAGGGCAGGAGAGTACCTGGGAATCATTACG
ATCCCTAGCAGCTCATCCTGCCCTTTGAATACCCTCACTTTCCAGGCCTGGCTCAGAATCTA
ACCTATTTTATTGACTGTCTGAGGGCCTTGAAAACAGGCCGAACCTGGAGGGCCTGGATTTT
TTTTTGGGCTGGAATGCTGCCCTGAGGGTGGGGCTGGCTCTTACTCAGGAAACTGCTGTGCC
CAACCCATGGACAGGCCAGCTGGGGCCACATGCTGACACAGACTCACTCAGAGACCCTTA
GACACTGGACCAGGCCTCCTCTCAGCCTTCTCTTTGTCCAGATTTCCAAAGCTGGATAAGTT
GGTCATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 54

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA61185

><subunit 1 of 1, 660 aa, 1 stop

><MW: 75220, pI: 6.76, NX(S/T): 0

MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQALRRFCQTGAVLFLLVTVIVNIKLILDTR
 RAISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEA
 REQGRGIHVIVLNQATGHVMAKRVFDTYSPHEDEAMVLFNLMVAPGRVLICTVKDEGSFHLK
 DTAKALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKHSKSPALSSWGDPVLLKTDVPLSS
 AEEAECHWADTELNRNRRRRFC SKVEGYGSVCCKDPTPIEFSPDPLPDNKVLNVPVAVIAGN
 RPNYLYRMLRSLLSAQGVSPQMITVFIDGYEPMDDVVALFGLRGIQHTPISIKNARVSQHY
 KASLTATFNLFPEAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSLYCISAWNDAQYEHTA
 EDPALLYRVETMPGLGWVLRSLYKEELEPKWPTPEKLWDWDMWMMRMPEQRRGRECIIPDVS
 RSYHFGIVGLNMNGYFHEAYFKKKHKFNTVPGVQLRNVDLSLKKEAYEVEVHRLLEAEVLDHS
 KNPCEDSFDPDTEGHTYVAFIRMEKDDDDFTTWTQLAKCLHIWDLDVIRGNHRGLWRLFRKKNH
 FLVVGVPASPYSVKKPPSVTPIFLEPPPKEEGAPGAPEQT

Important features of the protein:

Transmembrane domain:

amino acids 38-55

Homologous region to Mouse GNT1

amino acids 229-660

FIGURE 55

CGGACGCGTGGGCTGCTGGTGGGAAGGCCTAAAGAACTGGAAAGCCCACTCTCTTGGAAACCACCACAC
 CTGTTTAAAGAACCTAAGCACCATTAAAGCCCACTGGAAATTTGTTGTCTAGTGGTTGTGGGTGAATA
 AAGGAGGGCAGAAATGGGATGATTTTCATCTCCATTAGCCTGCTGTCTCTGGCTATGTTGGTGGGATGTTA
 CGTGGCCGGAATCATTTCCCTTGGCTGTTAATTTCTCAGAGGAACGACTGAAGCTGGTGAAGTGTGTTTGG
 GTGCTGGCCTTCTCTGTGGAAGTCTCTGGCAGTCACTGTCCTGAAGGAGTACATGCCCTTTATGAA
 GATATTCTTGAGGGAAAACACCACCAAGCAAGTGAAACACATAATGTGATTGCATCAGACAAAGCAGC
 AGAAAAATCAGTTGTCCATGAACATGAGCACAGCCACGACCACACACAGCTGCATGCCTATATTGGTG
 TTTCCCTCGTTCTGGGCTTCGTTTTTCATGTTGCTGGTGGACCAGATTGGTAACTCCCATGTGCATTCT
 ACTGACGATCCAGAAGCAGCAAGGTCTAGCAATTCCAAAATCACCACCACGCTGGGTCTGGTTGTCCA
 TGCTGCAGCTGATGGTGTGCTTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTTAATTG
 TGTTTGTGGCAATCATGCTACATAAGGCACCAGCTGCTTTTGGACTGGTTTCCTTCTTGATGCATGCT
 GGCTTAGAGCGGAATCGAATCAGAAAGCACTTGCTGGTCTTTGCATTGGCAGCACCAGTTATGTCCAT
 GGTGACATACTTAGGACTGAGTAAGAGCAGTAAAGAAGCCCTTTTTCAGAGGTGAACGCCACGGGAGTGG
 CCATGCTTTTCTCTGCCGGGACATTTCTTTATGTTGCCACAGTACATGTCTCCCTGAGGTGGGCGGA
 ATAGGGCACAGCCACAAGCCCGATGCCACGGGAGGGAGAGGCCTCAGCCGCCCTGGAAGTGGCAGCCCT
 GGTTCTGGGTTGCCTCATCCCTCTCATCCTGTGCTAGTAGGACACCAGCATTAATGTTCAAGGTCCAGC
 CTTGGTCCAGGGCCGTTTGCCATCCAGTGAGAACAGCCGGCAGCTGACAGCTACTCACTTCCTCAGTC
 TCTTGTCTCACCTTGCGCATCTCTACATGTATTCTTAGAGTCCAGAGGGGAGGTGAGGTTAAACCTG
 AGTAATGGAAAAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTATAGACATCCCATTGTGT
 TATCTTTTAAAGGCCCTTGACATTTTTCGTTTTAATATTCTCTTAACCTATTCTCAGGGAAGATG
 GAATTTAGTTTTAAGGAAAAGAGGAGAACTTCATACCAATGAAATAGTATTATGAAAATACAGT
 GTTCTGTAATTAAGCTATGTCTCTTTCTTCTTAGTTTTAGAGGCTCTGCTACTTTATCCATTGATTTTT
 AACATGGTTCACCACATGTAAGACTGGTGTCTTAGCATCTATGCCACATGCGTTGATGGAAGGTCATA
 GCACCCACTCACTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAATGATAGC
 AAGACACATTGAAAGCTCTCTTTTATACTCAAAGAGATATCCATTGAAAAGGGATGTCTAGAGGGATT
 TAAACAGCTCCTTTGGCACGTGCCTCTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGG
 TGGGAGGAGCTTCTAAAGAGGTGACTGGTATTTTGTAGCATTCCTTGTCAAGTTCTCCTTTGCAGAAT
 ACCTGTCTCCACATTCTAGAGAGGAGCCAAGTTCTAGTAGTTTCAGTTCTAGGCTTTTCTTCAAGAA
 CAGTCAGATCACAAGTGTCTTTGGAAATTAAGGGATATTAAATTTTAAAGTGATTTTTTGGATGGTTAT
 TGATATCTTTGTAGTAGCTTTTTTTTAAAGACTACCAAATGTATGGTTGTCCTTTTTTTTTTGT
 TTTTTTTTTTAATTATTTCTCTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGGTCAGCTTT
 GCGCAGACTGTGCTTCTCACATAACCACCTGTAGCAAGATGGATCATAAATGAGAAGTGTGTCCTA
 TTGATTTAAAGCTTATTGGAATCATGTCTCTTGTCTCTTCTCTTTCTTTGCTTTTCTTCTAACTTT
 TCCCTCTAGCCTCTCCTCGCCACAATTTGCTGCTTACTGCTGGTGTAAATATTGTGTGGGATGAATT
 CTTATCAGGACAACCCTTCTCGAAGTGAATAATGAAGATAATAATATCTTTATTCTTTATCCCCTT
 CAAAGAAATTACCTTTGTGTCAAATGCCGCTTTGTTGAGCCCTTAAATAACCACCTCCTCATGTGTAA
 ATTGACACAATCACTAATCTGGTAATTTAAACAATTGAGATAGCAAAGTGTTAACAGACTAGGATA
 ATTTTTTTTTTTCATATTGCCCCAAATTTTGTAAACCCTGTCTGTCAAATAAGTGTATAATATTGTAT
 TATTAATTTATTTTACTTTCTATACCATTTCAAACACATTACCTAAGGGGGAACCAAGACTAGTT
 TCTTCAGGGCAGTGGACGTAGTAGTTTGTAAAAACGTTTTCTATGACGCATAAGCTAGCATGCCTATG
 ATTTATTTCTTCATGAATTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTTGTGAGCCCTCTGCT
 GGCCACAGTGAGGAAAGTAGCACAAATAGGATACAGTTGTATGTAGTCATTGGCAACAATTGCATACA
 ATTTTACTACCAAGAGAAGGTATAGTATGGAAAGTCCAAATGACTTCCTTGATTGGATGTTAACAGCT
 GACTGGTGTGAGACTTGAGGTTTCATCTAGTCCTTCAAACCTATATGGTTGCCTAGATTCTCTCTGGA
 AACTGACTTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAAA

FIGURE 56

MDDFISISLLSLAMLVGCVAGIIPLAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHAL
YEDILEGKHHQASETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQ
IGNSHVHSTDDPEAARSSNSKITTTLGLVVHAAADGVALGAAASTSQTSVQLIVFVAIMLHK
APAAFGLVSFLMHAGLERNRIRKHLVLFALAAPVMSMVTYLGLSKSSKEALSEVNATGVAML
FSAGTFLYVATVHVLPEVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIPLILSVGHQH

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 37-56, 106-122, 211-230, 240-260, 288-304

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 57

GCTCGAGGCCGGCGGGCGGGAGAGCGACCCGGGCGGCCTCGTAGCGGGGCCCCGGATCCC
CGAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAGATGATGGGCTTGG
GAAACGGGCGTTCGACGATGAAGTCGCCGCCCTCGTGCTGGCCGCCCTGGTGGCCTGCATC
ATCGTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGACCTCCAGACACGGAT
CATGGAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGGCGCCGTGGAGCTGAAGA
AGAACGAGTTCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGC
CACAACCTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTTGGTGAA
TAACATCACACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGA
GGAATTACGGCAGGCTGCAGCAGGATGTCCTCCAGTTTCAGAAGAACCAGACCAACCTGGAG
AGGAAGTTCTCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTG
TGAGGAGCGAATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGA
GTGAAAACAACGACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAG
GCAGCAGGCCTGCCACACACAGAGGTGCCACAAGGGAAGGGAAACGTGCTTGGTAACAGCAA
GTCCCAGACACCAGCCCCCAGTTCCGAAGTGGTTTTTGGATTCAAAGAGACAAGTTGAGAAAG
AGGAAACCAATGAGATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAG
CCAGGCCGGGAGCAGGTGGTGGAAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGG
AGAACTGGGCCAGACCCCCACAGGTGCAGGCTGCCCTGTCAGTGAGCCAGGAAAATCCAGAGA
TGGAGGGCCCTGAGCGAGACCAGCTTGTTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCT
GCCGGGGAAGGGAGAAACCAGCAGAACTGAGAGGAGAAGATGACTACAACATGGATGAAAA
TGAAGCAGAATCTGAGACAGACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAAACATAGATG
TTTTTAATGTTGAAGATCAGAAAAGAGACACCATAAATTTACTTGATCAGCGTGAAAAGCGG
AATCATACACTCTTGAATTGAACTGGAATCACATATTTCAACAGGGCCGAAGAGATGACTA
TAAATGTTTCATGAGGGACTGAATACTGAAAACGTGTAATGTACTAAATAAAATGTACATCTGA

FIGURE 58

MMGLGNGRRSMKSPPLVLAALVACIIIVLGFNYWIIASSRSVDLQTRIMELEGRVRRAAAERGA
VELKKNEFQGELEKQREQLDKIQSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQL
KTLQRNYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVA
SRDLSENNDQRQQQLQALSEPQPRQLQAAGLPHTTEVPQGKGNVLGNSKSQTPAPSSEVVLDSCR
QVEKEETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQ
ENPEMEGPERDQLVIPDGQEEEQEAAGEGRNQQKLRGEDDYNMDENEAESETDKQAALAGND
RNIDVFNVEDQKRDTINLLDQREKRNHTL

Signal peptide:

amino acids 1-29

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 59

GGATGAGAGAAAGCCTCAGTGTGCTCTTCTGGCCTGGGTCTGCTTCTCTTCTACGCTGGCATTGCCCTCTTCA
 CCAGTGGCTTCTGCTCACCCGTTTGGAGCTCACCAACCATAGCAGCTGCCAAGAGCCCCAGGCCCTGGGTCCC
 TGCCATGGGGGAGCCAAGGGAAACCTGGGGCTGCTGGATGGCTTCCGATTTTCGCGGGTTGTGTTGGTGCTGA
 TAGATGCTCTGCGATTTGACTTCGCCCAGCCCCAGCATTACACGTGCCTAGAGAGCCTCCTGTCTCCCTACCCT
 TCCTGGGCAAACCTAAGCTCCTTGCAGAGGATCCTGGAGATTACAGCCCCACCATGCCCGGCTCTACCGATCTCAGG
 TTGACCCTCCTACCACCACCATGCAGCGCCTCAAGGCCCTCACCCTGGCTCACTGCCTACCTTTATTGATGCTG
 GTAGTAACCTCGCCAGCCACGCCATAGTGAAGACAATCTCATTAAGCAGCTCACCAGTGCAGGAAGGCGTGTAG
 TCTTCATGGGAGATGATACCTGGAAAGACCTTTTCCCTGGTGCTTCTCCAAAGCTTTCTTCTTCCCATCCTTCA
 ATGTCAGAGACCTAGACACAGTGGACAATGGCATCCTGGAAACACCTCTACCCACCATGGACAGTGGTGAATGGG
 ACGTGCTGATTGCTCACTTCTGGGTGTGGACCACTGTGGCCACAAGCATGGCCCTCACCACCCTGAAATGGCCA
 AGAACTTAGCCAGATGGACCAGGTGATCCAGGGACTTGTGGAGCGTCTGGAGAATGACACACTGCTGGTAGTGG
 CTGGGGACCATGGGATGACCACAAATGGAGACCATGGAGGGGACAGTGAGCTGGAGGTCTCAGCTGCTCTCTTTC
 TGTATAGCCCCACAGCAGTCTTCCCCAGCACCCCCACAGAGGAGCCAGAGGTGATTCTCAAGTTAGCCTTGTGC
 CCACGCTGGCCCTGCTGCTGGGCCTGCCATCCCATTGGGAATATCGGGGAAGTGATGGCTGAGCTATTCTCAG
 GGGGTGAGGACTCCAGCCCCACTCCTCTGCTTTAGCCCAAGCCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT
 CCGGATTTCTTCATACCTACTCAGCTGCTACTCAGGACCTTCAAGCTAAGGAGCTTCATCAGCTGCAGAACCTCT
 TCTCCAAGGCCTCTGCTGACTACCAGTGGCTTCTCCAGAGCCCCAAGGGGGCTGAGGCGACACTGCCGACTGTGA
 TTGCTGAGCTGCAGCAGTTCCTGCGGGGAGCTCGGGCCATGTGCATCGAGTCTTGGGCTCGTTTCTCTCTGGTCC
 GCATGGCGGGGGTACTGCTCTCTTGGCTGCTTCTGCTTTATCTGCCTGCTGGCATCTCAGTGGGCAATATCCC
 CAGGCTTTCCATTCTGCCCTCTACTCCTGACACCTGTGGCCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC
 TCCTGGGAACCTATTGAGCTGAAGCTAGATCTAGTGCTTCTAGGGGCTGTGGCTGCAGTGAGCTCATTCTCCCTT
 TTCTGTGGAAGCCTGGGCTGGCTGGGGGTCCAGAGGCCCTGGCAACCCTGTTTCCCATCCCTGGGCCCGTCC
 TGTTACTCCTGCTGTTTCGCTTGGCTGTGTTCTTCTCTGATAGTTTGTGTAGCTGAGGCCAGGGCCACCCCT
 TCCTTTTGGGCTCATTCTCCTGCTCCTGGTTGTCCAGCTTCACTGGGAGGGCCAGCTGCTTCCACCTAAGCTAC
 TCACAATGCCCCGCCTTGGCACTTCAGCCACAACAAACCCCCACGGCACAATGGTGCATATGCCCTGAGGCTTG
 GAATTGGGTGCTTTTATGTACAAGGCTAGCTGGGCTTTTTCATCGTTGCCCTGAAGAGACACCTGTTTGCCACT
 CCTCTCCCTGGCTGAGTCCCTCTGGCATCCATGGTGGTGGTTCGAGCCAAGAATTTATGGTATGGAGCTTGTGTGG
 CGGCGCTGGTGGCCCTGTTAGCTGCCGTGCGCTTGTGGCTTCGCCGCTATGGTAATCTCAAGAGCCCCGAGCCAC
 CCATGCTCTTTGTGCGTGGGGACTGCCCTAATGGCATTGGGTACTGCTGCCTACTGGGCATTGGCGTCGGGGG
 CAGATGAGGCTCCCCCGCTCTCCGGTCTGCTCTCTGGGGCATCCATGGTGTGCTCGGCTGTAGCAGGGC
 TGGCTGCTTCAGGGCTCGCGCTGCTGCTCTGGAAGCCTGTGACAGTGCTGGTGAAGGCTGGGGCAGGCGCTCCAA
 GGACCAGGACTGCTCACTCCCTTCTCAGGCCCCCCTTCTCAAGCTGACTTGGATTATGTGGTCCCTCAA
 TCTACCGACACATGCAGGAGGAGTTCGGGGCCGTTAGAGAGGACCAATCTCAGGGTCCCCTGACTGTGGCTG
 CTTATCAGTTGGGGAGTGTTACTCAGCTGCTATGGTACAGCCCTCACCCTGTTGGCCTTCCCCTTCTGCTGT
 TGCATGCGGAGCGCATCAGCCTTGTGTTCTGCTTCTGTTTCTGCAGAGCTTCTTCTCCTACATCTGCTTGTG
 CTGGGATACCCGTCACCACCCTGGTCTTTTACTGTGCCATGGCAGGCACTCTCGGCTTGGGCCCTCATGGCCA
 CACAGACCTTCTACTCCACAGGCCACCAGCCTGTCTTCCAGCCATCCATTGGCATGCAGCCTTCTGGGATTCC
 CAGAGGTCATGGCTCCTGTACTTGGCTGCCTGCTTGTCTAGTGGGAGCCAACACCTTGGCTCCCACCTCCTCT
 TTGCAGTAGGTTGCCCACTGCTCCTGCTCTGGCCTTCTGTTGTGAGAGTCAAGGGCTGCGGAAGAGACAGCAGC
 CCCCAGGGAATGAAGCTGATGCCAGAGTCAAGCCGAGGAGGAAGAGGAGCCACTGATGGAGATGCGGCTCCGGG
 ATGCGCCTCAGCACTTCTATGCAGCACTGCTGCAGCTGGGCCTCAAGTACCTCTTTATCCTTGGTATTAGATT
 TGGCCTGTGCTTGGCAGCCTCCATCCTTCGAGGCATCTCATGGTCTGGAAAGTGTGTTGCCCTAAGTTCATAT
 TTGAGGCTGTGGGCTTCAATGTGAGCAGCGTGGGACTTCTCCTGGGCATAGCTTGGTGTGAGAGTGGATGGTG
 CTGTGAGCTCCTGGTTCAGGCAGCTATTTCTGGCCCAGCAGAGGTAGCCTAGTCTGTGATTACTGGCACTTGGCT
 ACAGAGAGTGCTGGAGAACAGTGTAGCCTGGCCTGTACAGGTACTGGATGATCTGCAAGACAGGCTCAGCCATAC
 TCTTACTATCATGCAGCCAGGGGCCGCTGACATCTAGGACTTCAATTATTCTATAATTACAGGACCACAGTGGAGTA
 TGATCCCTAACTCCTGATTTGGATGCATCTGAGGGACAAGGGGGGCGGTCTCCGAAGTGAATAAAATAGGCCGG
 GCGTGGTGACTTGACCTATAATCCAGCACTTTGGGAGGCAGAGGTGGGAGGATTGCTTGGTCCCAGGAGTTCA
 AGACCAGCCTGTGGAACATAACAAGACCCCGTCTCTACTATTTAAAAAAGTGTAAATAAATGATAATAT

FIGURE 60

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809

<subunit 1 of 1, 1089 aa, 1 stop

<MW: 118699, pI: 8.49, NX(S/T): 2

MQKASVLLFLAWVCFLFYAGIALFTSGFLLTRLELTNHSSCQEPPGPGSLPWGSQGKPGACW
 MASRFSRVVLVLIDALRFDAQPQSHVPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQ
 VDPPTTTMQRLLKALTTGSLPTFIDAGSNFASHAIVEDNLIKQLTSAGRRVVFMDGDDTWKDLF
 PGAFSKAFFFPFSFNVRDLDTVDNGILEHLYPTMDSGEWDVLIHFLGVDHCGHKHGHHPHEM
 AKKLSQMDQVIQGLVERLENDTLLVVAGDHGTTNGDHGGDSELEVSAALFLYSPTAVFPST
 PPEEPEVIPQVSLVPTLALLLGLPIPFNGIGEVMAELEFSGGEDSQPHSSALAQASALHLNAQ
 QVSRFLHTYSAATQDLQAKELHQLQNLFSKASADYQWLLQSPKGAEATLPTVIAELQQFLRG
 ARAMCIESWARFSLVRMAGGTALLAASCFICLLASQWAI SPGFPPCPLLLTPVAWGLVGAIA
 YAGLLGTIELKLDLVLLGAVAAVSSFLPFLWKAWAGWGSKRPLATLFPPIPGPVLLLLLLFRLA
 VFFSDSFVVAEARATPFLLGSFILLLVVQLHWEGQLLPKLLTMPRLGTSATTNPPRHNGAY
 ALRLGIGLLLCTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVAALVALLA
 AVRLWLRRYGNLKSPEPPMLFVRWGLPLMALGTAAWALASGADEAPPRLRVLVSGASMVLP
 RAVAGLAASGLALLLWKPVTVLVKAGAGAPRTRTVLTPFSGPPTSQADLDYVVPQIYRHMQE
 EFRGRLERTKSQGPLTVAAYQLGSVYSAMVTALTLLAFPLLLLHAERISLVFLLLFLQSFL
 LLHLLAAGIPVTTGPGFTVPWQAVSAWALMATQTFYSTGHQPVFPAIHWHAAFVGVFPEGHGS
 CTWLPALLVGANTFASHLLFAVGCPLLLLWPFLLCESQGLRKRQQPPGNEADARVRPEEEEEEP
 LMEMRLRDAPQHFFYAALLQLGLKYLFI LGIQILACALAASILRRHLMVWKVFAPKFI FEAVG
 FIVSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domains:

amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850,
 1016-1034, 1052-1070

Leucine zipper pattern.

amino acids 843-864

N-glycosylation sites.

amino acids 37-40, 268-271

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 61

[illegible]

FIGURE 62

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62815

<subunit 1 of 1, 442 aa, 1 stop

<MW: 49932, pI: 4.55, NX(S/T): 5

MSYNGLHQRVFKELKLLTLCSISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQ
QIYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQ
CARTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDF
KRFFVPAEKIVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKH
LGYASHLMEIFCDSEENTEGTSLTQQESLSRTIPDKTVIEYDYDVRTTDICAGPEEQELSL
QEEVSTQGTLLLESQAALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEEPSTTLVDWD
PQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEW
GLYVQMEN

Important features:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 63

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCG
 TCTGCC**ATG**GGGCTCGGGTTGAGGGGCTGGGGACGTCCTCTGCTGACTGTGGCCACCGCCCT
 GATGCTGCCCCGTGAAGCCCCCGCAGGCTCCTGGGGGGCCAGATCATCGGGGGCCACGAGG
 TGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGGCCAACATCACTGCGGA
 GGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGACCT
 CCGCACTGGCCTGGTGGTGTGGGCGCCACGTCCTGAGTACTGCGGAGCCCACCCAGCAGG
 TGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGAC
 ATCTGCCCTGCTGCGGCTGAACGGCTCTGCTGTCCTGGGCCCTGCAGTGGGGCTGCTGAGGCT
 GCCAGGGAGAAGGGCCAGGCCCCCACAGCGGGGACACGGTGCCGGGTGGCTGGCTGGGGCT
 TCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCCAAGGTCCGAGTGCTGGAC
 CCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCACCCGCAG
 TGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGCCCCTGGTGTGCAGGA
 ACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCGAC
 GTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGACGTGGTTCGGCGGAGCAGTCC
 CCAGCCCGGCCCCCTGCCTGGGACCACCAGGCCCCAGGAGAAGCCGCC**TGAG**CCACAACCT
 TCGGGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCGTGGCTGGGCCCCACGGG
 AAGCCTGATGTTCAAGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACATGCA
 AAGGGCAGAAGCAAACCCAGTAAAATGTTAACTGACAAAAAAAAAAAAAAAAAAAAAGAAA

FIGURE 65

GAGCTACCCAGGCGGCTGGTGTGCAGCAAGCTCCGCGCCGACTCCGGACGCCTGACGCCTGA
 CGCCTGTCCCCGGCCCCGGC**ATG**AGCCGCTACCTGCTGCCGCTGTCGGCGCTGGGCACGGTAG
 CAGGCGCCGCCGTGCTGCTCAAGGACTATGTCACCGGTGGGGCTTGCCCCAGCAAGGCCACC
 ATCCCTGGGAAGACGGTCATCGTGACGGGCGCCAACACAGGCATCGGGAAGCAGACCGCCTT
 GGAAGTGGCCAGGAGAGGAGGCAACATCATCTGGCCTGCCGAGACATGGAGAAGTGTGAGG
 CGGCAGCAAAGGACATCCGCGGGGAGACCCTCAATCACCATGTCAACGCCCCGGCACCTGGAC
 TTGGCTTCCCTCAAGTCTATCCGAGAGTTTGCAGCAAAGATCATTTGAAGAGGAGGAGCGAGT
 GGACATTCTAATCAACAACGCGGGTGTGATGCGGTGCCCCACTGGACCACCGAGGACGGCT
 TCGAGATGCAGTTTGGCGTTAACCACCTGGGTCACTTTCTCTTGACAACTTGCTGCTGGAC
 AAGCTGAAAGCCTCAGCCCCCTTCGCGGATCATCAACCTCTCGTCCCTGGCCCATGTTGCTGG
 GCACATAGACTTTGACGACTTGAAGTGGCAGACGAGGAAGTATAACACCAAAGCCGCCTACT
 GCCAGAGCAAGCTCGCCATCGTCCTCTTACCAAGGAGCTGAGCCGGCGGCTGCAAGGCTCT
 GGTGTGACTGTCAACGCCCTGCACCCCGGCGTGGCCAGGACAGAGCTGGGCAGACACACGGG
 CATCCATGGCTCCACCTTCTCCAGCACCACACTCGGGCCCATCTTCTGGCTGCTGGTCAAGA
 GCCCCGAGCTGGCCGCCAGCCCAGCACATACTGGCCGTGGCGGAGGAAGTGGCGGATGTT
 TCCGGAAAGTACTTCGATGGACTCAAACAGAAGGCCCCGGCCCCCGAGGCTGAGGATGAGGA
 GGTGGCCCCGAGGCTTTGGGCTGAAAGTGCCCGCCTGGTGGGCTTAGAGGCTCCCTCTGTGA
 GGGAGCAGCCCCCTCCCCAGAT**TAA**CCTCTGGAGCAGATTTGAAAGCCAGGATGGCGCCTCCAG
 ACCGAGGACAGCTGTCCGCCATGCCCGCAGCTTTCCTGGCACTACCTGAGCCGGGAGACCCAG
 GACTGGCGGCCGCCATGCCCGCAGTAGGTTCTAGGGGGCGGTGCTGGCCGCAGTGGACTGGC
 CTGCAGGTGAGCACTGCCCCGGGCTCTGGCTGGTTCCGTCTGCTCTGCTGCCAGCAGGGGAG
 AGGGGCCATCTGATGCTTCCCCTGGGAATCTAAACTGGGAATGGCCGAGGAGGAAGGGGCTC
 TGTGCACTTGCAGGCCACGTGAGGAGAGCCAGCGGTGCCTGTGCGGGAGGGTTCCAAGGTGC
 TCCGTGAAGAGCATGGGCAAGTTGTCTGACACTTGGTGGATTCTTGGGTCCCTGTGGGACCT
 TGTGCATGCATGGTCTCTCTGAGCCTTGGTTTCTTACGAGTGAGATGCTCAGAATAACTG
 CTGTCTCCCATGATGGTGTGGTACAGCGAGCTGTTGTCTGGCTATGGCATGGCTGTGCCGGG
 GGTGTTTGTGAGGGCTTCCCTGTGCCAGAGCCCAGCCAGAGAGCAGGTGCAGGTGTCATCCC
 GAGTTCAGGCTCTGCACGGCATGGAGTGGGAACCCACACAGCTGCTGCTACAGGACCTGGGA
 TTGCCTGGGACTCCCACCTTCCATCAATTCTCATGGTAGTCCAAACTGCAGACTCTCAAAC
 TTGCTCATTT

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64842

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35932, pI: 8.45, NX(S/T): 1

MSRYLLPLSALGTVAGAAVLLKDYVTGGACPSKATIPGKTVIVTGANTGIGKQTALELARRG
GNIILACRDMEKCEAAAKDIRGETLNHHVNARHLDLASLKSIREFAAKIIEEEEERVDILINN
AGVMRCPHWTTEDGFEMQFGVNH LGHFLLTNLLLDK LKASAPSRIINLSSLAHVAGHIDFDD
LNWQTRKYNTKAAYCQSKLAIVLFTKELSRRLQGSGVTVNALHPGVARTELGRHTGIHGSTF
SSTTLGP IFWLLVKSP ELAAQPSTYLAVAEELADVSGKYFDGLKQKAPAEAEDEEVARRLW
AESARLVGLEAPSVREQPLPR

Signal peptide:

amino acids 1-17

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 67

GAAGTTCGCGAGCGCTGGC**ATG**TGGTCCTGGGGCGCGGCTGGCGGCGCTGCTGGCGGTGCTG
 GCGCTCGGGACAGGAGACCCAGAAAGGGCTGCGGCTCGGGGCGACACGTTCTCGGCGCTGAC
 CAGCGTGGCGCGCGCCCTGGCGCCCGAGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGC
 GCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCTACGACAAGGTACTTTCTTTGCAT
 GAGGATTCAACAACCCCTGTGGCTAACCCTCTGCTTGCATTTACTCTCATCAAACGCCTGCA
 GTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGG
 ATGGCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGG
 GCCCTGATGCGGCTGCAGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCCGAGGTGTCTT
 TCAGAGAGTCACTGGCTCTGCCATCACTGACCTGTACAGCCCCAAACGGCTCTTTTCTCTCA
 CAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCTATGACATGGGGGATTATTACCATGCC
 ATTCCATGGCTGGAGGAGGCTGTCAGTCTCTTCCGAGGATCTTACGGAGAGTGGAAGACAGA
 GGATGAGGCAAGTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCAGGAA
 ATGTTTCGTGTGCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCCAGATAATAAGAGG
 ATGGCCAGGAATGTCTTGAAATATGAAAGGCTCTTGGCAGAGAGCCCCAACACGTGGTAGC
 TGAGGCTGTCATCCAGAGGCCCAATATACCCACCTGCAGACCAGAGACACCTACGAGGGGC
 TATGTCAGACCCTGGGTTCCCAGCCCCTCTCTACCAGATCCCTAGCCTCTACTGTTCCCTAT
 GAGACCAATTCCAACGCCTACCTGCTGCTCCAGCCCATCCGGAAGGAGGTCTACACCTGGA
 GCCCTACATTGCTCTCTACCATGACTTCGTCTAGTCACTCAGAGGCTCAGAAAATTAGAGAAC
 TTGCAGAACCATGGCTACAGAGGTCACTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAG
 TACCGCATCAGCAAAAGTGCTGGCTGAAGGACACTGTTGACCCAAAACCTGGTGACCCTCAA
 CCACCGCATTGCTGCCCTCACAGGCCTTGATGTCCGGCCTCCCTATGCAGAGTATCTGCAGG
 TGGTGAACATATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATGCTACGTACCAAGC
 AGCCCCCTCTACAGAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCTC
 GGTGGAAGCTGGAGGAGCCACAGCCTTCATCTATGCCAACCTCAGCGTGCCTGTGGTTAGGA
 ATGCAGCACTGTTTTGGTGGAACTGCACAGGAGTGGTGAAGGGGACAGTGACACACTTCAT
 GCTGGCTGTCCTGTCCTGGTGGGAGATAAGTGGGTGGCCAACAAGTGGATACATGAGTATGG
 ACAGGAATTCCGCAGACCCTGCAGCTCCAGCCCTGAAGAC**TGA**ACTGTTGGCAGAGAGAAGC
 TGGTGGAGTCCGTGTGGCTTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTAGGAGAGGAGAA
 AGCAGAGCAGCCTCCTGGAAGAAGGCCTTGTCAGCTTTGTCTGTGCCTCGCAAATCAGAGGC
 AAGGGAGAGGTTGTTACCAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAA
 GTCAGAGTAGGATGCACAGTACAAAGGAGGGGGGAGTGGAGGCCTGAGAGGGAAGTTTCTGG
 AGTTCAGATACTCTCTGTTGGGAACAGGACATCTCAACAGTCTCAGGTTTCGATCAGTGGGTC
 TTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGGCAATGAGGACACCTGCAGGAG
 GGGCTAGCCTGACTCCCAGAACTTTAAGACTTTCTCCCCACTGCCTTCTGCTGCAGCCCAAG
 CAGGGAGTGTCCCCCTCCCAGAAAGCATATCCCAGATGAGTGGTACATTATATAAGGATTTTT
 TTTAAGTTGAAAACAACCTTTCTTTTCTTTTGTATGATGGTTTTTTAACACAGTCATTAAAA
 ATGTTTATAAATCAAAA

FIGURE 68

MGP GARLAALLAVLALGTGDPERAAARGDTFSALTSVARALAPERRLLGLLRRYLRGEEARL
RDLTRFYDKVLSLHEDSTTPVANPLLAFTTLIKRLQSDWRNVVHSLEASENIRALKDGYEKVE
QDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRVTGSAITDLYSPKRLFSLTGDDCFQ
VGKVAYDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRAGNVSCALS
LSREFLLYSPDNKRMARNVLKYERLLAESPNHVVAEAVIQRPNI PHLQTRDTYEGLCQTLGS
QPTLYQIPSLYCSYETNSNAYLLLQPIRKEVIHLEPYIALYHDFVSDSEAQKIRELAEPWLQ
RSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRPPYAEYLQVVNYGIG
GHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLLSSVEAGGATAFIYANLSVPVVRNAALFWW
NLHRSGEGDSDTLHAGCPVLVGDKWVANKWIHEYGQEFRRPCSSSPED

Signal peptide:

amino acids 1-19

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 69

GAGATAGGGAGTCTGGGTTTAAGTTCCTGCTCCATCTCAGGAGCCCCCTGCTCCCACCCCTAG
 GAAGCCACCAGACTCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCC
 ACGAGCGCTGGCTGAGGGACCGAGCCGGAGAGCCCCGGAGCCCCCGTAACCCGCGCGGGGAG
 CGCCCAGGATGCCGCGCGGGGACTCGGAGCAGGTGCGCTACTGCGCGCGCTTCTCCTACCTC
 TGGCTCAAGTTTTCACTTATCATCTATTCCACCGTGTCTGGCTGATTGGGGCCCTGGTCCT
 GTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCCTTGAAAGTGCCTTCC
 TGGCTCCAGCCATCATCCTCATCCTCCTGGGCGTCGTCATGTTTCATGGTCTCCTTCATTGGT
 GTGCTGGCGTCCCTCCGTGACAACCTGTACCTTCTCCAAGCATTTCATGTACATCCTTGGGAT
 CTGCCTCATCATGGAGCTCATTGGTGGCGTGGTGGCCTTGACCTTCCGGAACCAGACCATTG
 ACTTCCTGAACGACAACATTCGAAGAGGAATTGAGAACTACTATGATGATCTGGACTTCAAA
 AACATCATGGACTTTGTTTCAGAAAAAGTTCAAGTGCTGTGGCGGGGAGGACTACCGAGATTG
 GAGCAAGAATCAGTACCACGACTGCAGTGCCCCCTGGACCCCTGGCCTGTGGGGTGCCCTACA
 CCTGCTGCATCAGGAACACGACAGAAGTTGTCAACACCATGTGTGGCTACAAAACCTATCGAC
 AAGGAGCGTTTCAGTGTGCAGGATGTCATCTACGTGCGGGGCTGCACCAACGCCGTGATCAT
 CTGGTTTCATGGACAACCTACACCATCATGGCGTGCATCCTCCTGGGCATCCTGCTTCCCCAGT
 TCCTGGGGGTGCTGCTGACGCTGCTGTACATCACCCGGGTGGAGGACATCATCATGGAGCAC
 TCTGTCACTGATGGGCTCCTGGGGCCCGGTGCCAAGCCCAGCGTGGAGGCGGCAGGCACGGG
 ATGCTGCTTGTGCTACCCCAATTAGGGGCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTC
 TGGGATAGCACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCTCTGCCACACA
 CTCAGTACTGACCAAAGCCAGGGCTGTGTGTGCCTGTGTGTAGGTCCCACGGCCTCTGCCTC
 CCCAGGGAGCAGAGCCTGGGCCTCCCCTAAGAGGCTTTCCCCGAGGCAGCTCTGGAATCTGT
 GCCCACCTGGGGCCTGGGGAACAAGGCCCTCCTTTCTCCAGGCCTGGGCTACAGGGGAGGGA
 GAGCCTGAGGCTCTGCTCAGGGCCCATTTTCATCTCTGGCAGTGCCTTGGCGGTGGTATTCAA
 GGCAGTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCTCGGGGCAGGAGGGAAGG
 GCATCTGGGGAAGGGCAGGAGGGAAGAGCTGTCCATGCAGCCACGCCCATGGCCAGGTTGGC
 CTCTTCTCAGCCTCCCAGGTGCCTTGAGCCCTCTTGCAAGGGCGGCTGCTTCCTTGAGCCTA
 GTTTTTTTTTTACGTGATTTTTGTAAACATTCATTTTTTTTGTACAGATAACAGGAGTTTCTGAC
 TAATCAAAGCTGGTATTTCCCCGCATGTCTTATTCTTGCCCTTCCCCCAACCAGTTTGTTAA
 TCAAACAATAAAAAACATGTTTTGTTTTGTTTTTAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863

><subunit 1 of 1, 294 aa, 1 stop

><MW: 33211, pI: 5.35, NX(S/T): 3

MPRGDSEQVRYCARFSYLWLKFSLLIYSTVFWLIGALVLSVGIYAEVERQKYKTLESAFLAP
AIILILLGVVMFMVSFIGVLASLRDONLYLLQAFMYILGICLIMELIGGVVALTFRNQTIIDFL
NDNIRRGIEENYYDDLDFKNIMDFVQKKFKCCGGEDYRDWSKNQYHDCSAPGPLACGVPTCC
IRNTTEVVNTMCGYKTIDKERFSVQDVIYVRGCTNAVLIWFMDNYTIMACILLGILLPQFLG
VLLTLLYITRVEDIIMEHSVTDGLLGPGAKPSVEAAGTGCCCLCPN

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 22-42, 57-85, 93-116, 230-257

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 71

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGA
 CACCTGGGAAG**ATG**GCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACC
 TTGATCCAAGCCACCCTCAGTCCCAGTTCATCCTCGGCCCAAAGTCATCAAAGA
 AAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGC
 TCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACC
 GTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAA
 GCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCA
 ACACGCCCCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATC
 CGCATGGACACCAGTGCAAGTGGCCCCACCCGCCTGGTCCTCAGTGAAGTGTGCCACCAGCCA
 TGGGAGCCTGCGCATCCAAGTGTGTATAAGCTCTCCTTCCCTGGTGAACGCCTTAGCTAAGC
 AGGTCATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTG
 ATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTTC
 CCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTTC
 AGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAAC
 TCTGCAGCTTCCCTGACAATGCCCCACCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCA
 GGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTCATGGTCCTGTTGG
 ACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAG
 GCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTT
 TTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTCCCTCCA
 GTGAAGCCCTCCGCCCTTTGTTACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTAC
 ACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGAT
 GAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCC
 ACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTG
 AAGGCCTTGGGATTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCC
 AGCCTCCTTGTGGAAACCCAGCTCTCCTGTCTCCCAG**TGA**AGACTTGGATGGCAGCCATCAG
 GGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAAT
 CAATAAACACTTGCTGTGAAAAA

FIGURE 72

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881

><subunit 1 of 1, 484 aa, 1 stop

><MW: 52468, pI: 7.14, NX(S/T): 3

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAM
REKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPL
VKTIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMN
LLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITIQLYL
GAKLLDSQGKVTKWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVL
PESAHRLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEAL
RPLFTLGIEASSEAQFYTKGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSIL
LPNQNGKLRSGVPVSLVKALGFEEAESSLTkdALVLTpasLWKpSSPVSQ

Important features of the protein:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 73

GAGCGAAC**ATG**GCAGCGCGTTGGCGGTTTTGGTGTGTCTCTGTGACCATGGTGGTGGCGCTG
 CTCATCGTTTTGCGACGTTCCCTCAGCCTCTGCCCCAAAGAAAGAGGAGATGGTGTTATCTGA
 AAAGGTTAGTCAGCTGATGGAATGGACTAACAAAAGACCTGTAATAAGAATGAATGGAGACA
 AGTTCCGTCGCCTTGTGAAAGCCCCACCGAGAAATTAATCCGTTATCGTCATGTTCACTGCT
 CTCCAACCTGCATAGACAGTGTGTCGTTTGCAAGCAAGCTGATGAAGAATTCCAGATCCTGGC
 AAACCTCTGGCGATACTCCAGTGCATTCACCAACAGGATATTTTTTGCCATGGTGGATTTTG
 ATGAAGGCTCTGATGTATTTTCAAGATGCTAAACATGAATTCAGCTCCAACTTTCATCAACTTT
 CCTGCAAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGGTGCGGGGTTTTTTCAGC
 TGAGCAGATTGCCCGGTGGATCGCCGACAGAACTGATGTCAATATTAGAGTGATTAGACCCC
 CAAATTATGCTGGTCCCCTTATGTTGGGATTGCTTTTGGCTGTTATTGGTGGACTTGTGTAT
 CTTGCAAGAAGTAATATGGAATTTCTCTTTAATAAACTGGATGGGCTTTTGCAGCTTTGTG
 TTTTGTGCTTGCTATGACATCTGGTCAAATGTGGAACCATATAAGAGGACCACCATATGCCC
 ATAAGAATCCCCACACGGGACATGTGAATTATATCCATGGAAGCAGTCAAGCCCAGTTTGTA
 GCTGAAACACACATTGTTCTTCTGTTTAATGGTGGAGTTACCTTAGGAATGGTGGCTTTTATG
 TGAAGCTGCTACCTCTGACATGGATATTGGAAGCGAAAGATAATGTGTGTGGCTGGTATTG
 GACTTGTGTATTATTCTTCAGTTGGATGCTCTCTATTTTTAGATCTAAATATCATGGCTAC
 CCATACAGCTTTCTGATGAGT**TAA**AAAGGTCCCAGAGATATATAGACACTGGAGTACTGGAA
 ATTGAAAAACGAAAATCGTGTGTGTTTGAAAAAGAAGATGCAACTTGTATATTTTGTATTAC
 CTCTTTTTTTCAAGTGATTTAAATAGTTAATCATTTAACCAGAAAGATGTGTAGTGCCTTA
 ACAAGCAATCCTCTGTCAAAATCTGAGGTATTTGAAAATAATTATCCTCTTAACCTTCTCTT
 CCCAGTGAACCTTTATGGAACATTTAATTTAGTACAATTAAGTATATTATAAAAATTGTAAA
 CTACTACTTTGTTTTAGTTAGAACAAAGCTCAAACTACTTTAGTTAACTTGGTCATCTGAT
 TTTATATTGCCTTATCCAAAGATGGGGAAAGTAAGTCCTGACCAGGTGTTCCACATATGCC
 TGTTACAGATAACTACATTAGGAATTCATTCTTAGCTTCTTCATCTTTGTGTGGATGTGTAT
 ACTTTACGCATCTTTCCCTTTTGGAGTAGAGAAATTATGTGTGTCATGTGGTCTTCTGAAAATG
 GAACACCATTCTTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTTCTCCTCCTCCTT
 GCATATTTCCCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCTCAAAAAAAGTA
 TCTCTAAATACAGGATTATAATTTCTGCTTGAGTATGGTGTAACTACCTTGTATTTAGAAA
 GATTTTCAGATTCAATCCATCTCCTTAGTTTTCTTTTAAAGGTGACCCATCTGTGATAAAAATA
 TAGCTTAGTGCTAAAATCAGTGTAACCTTATACATGGCCTAAAATGTTTCTACAAATTAGAGT
 TTGTCACTTATTCCATTTGTACCTAAGAGAAAAATAGGCTCAGTTAGAAAAGGACTCCCTGG
 CCAGGCGCAGTGACTTACGCCTGTAATCTCAGCACTTTGGGAGGCCAAGGCAGGCAGATCAC
 GAGGTCAGGAGTTCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATAT
 AAAAATTAGCTGGGTGTGGTGGCAGGAGCCTGTAATCCCAGCTACACAGGAGGCTGAGGCAC
 GAGAATCACTTGAACCTCAGGAGATGGAGGTTTCAGTGAGCCGAGATCACGCCACTGCACTCC
 AGCCTGGCAACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAA

FIGURE 74

MAARWRFWCVSVTMVVALLIVCDVPSASAQRKKEMVLSEKVSQLMEWTKRPPVIRMNGDKFR
RLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIFAMVDFDEG
SDVFQMLNMNSAPTFINFPAKGKPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNY
AGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFALCFVLAMTSGQMWNHIRGPPYAHKN
PHTGHVNYIHGSSQAQFVAETHIVLLENGGVTLGMVLLCEAATSDMDIGKRKIMCVAGIGLV
VLFFSWMLSIERSKYHGYPYSFLMS

Signal peptide:

amino acids 1-29

Transmembrane domains:

amino acids 183-205, 217-237, 217-287, 301-321

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 75

AAGCAACCAAACCTGCAAGCTTTGGGAGTTGTTTCGCTGTCCCTGCCCTGCTCTGCTAGGGAGA
 GAACGCCAGAGGGAGGCGGCTGGCCCGGCGGCAGGCTCTCAGAACCGCTACCGGCG**ATG**CTA
 CTGCTGTGGGTGTCTGGTGGTCGCAGCCTTGGCGCTGGCGGTACTGGCCCCGGAGCAGGGGA
 GCAGAGGCGGAGAGCAGCCAAAGCGCCCAATGTGGTGCTGGTTCGTGAGCGACTCCTTCGATG
 GAAGGTAAACATTTTCATCCAGGAAGTCAGGTAGTGAAACTTCCTTTTATCAACTTTATGAAG
 ACACGTGGGACTTCCTTTCTGAATGCCTACACAAACTCTCCAATTTGTTGCCCATCACGCGC
 AGCAATGTGGAGTGGCCTCTTCACTCACTTAACAGAATCTTGGAATAATTTTAAGGGTCTAG
 ATCCAAATTATACAACATGGATGGATGTCATGGAGAGGCATGGCTACCGAACACAGAAATTT
 GGGAAACTGGACTATACTTCAGGACATCACTCCATTAGTAATCGTGTGGAAGCGTGGACAAG
 AGATGTTGCTTTCTTACTCAGACAAGAAGGCAGGCCCATGGTTAATCTTATCCGTAACAGGA
 CTAAAGTCAGAGTGATGGAAAGGGATTGGCAGAATACAGACAAAGCAGTAAACTGGTTAAGA
 AAGGAAGCAATTAATTACACTGAACCATTTGTTATTTACTTGGGATTAAATTTACCACACCC
 TTACCCTTCACCATCTTCTGGAGAAAATTTTGGATCTTCAACATTTACACATCTCTTTATT
 GGCTTGAAAAAGTGTCTCATGATGCCATCAAAATCCCAAAGTGGTCACCTTTGTCAGAAATG
 CACCCTGTAGATTATTACTCTTCTTATACAAAAAACTGCACTGGAAGATTTACAAAAAAGA
 AATTAAGAATATTAGAGCATTTTATTATGCTATGTGTGCTGAGACAGATGCCATGCTTGGTG
 AAATTATTTTGGCCCTTCATCAATTAGATCTTCTTCAGAAAACCTATTGTCATATACTCCTCA
 GACCATGGAGAGCTGGCCATGGAACATCGACAGTTTTTATAAAATGAGCATGTACGAGGCTAG
 TGCACATGTTCCGCTTTTGATGATGGGACCAGGAATTAAAGCCGGCCTACAAGTATCAAATG
 TGGTTTCTCTTGTGGATATTTACCCTACCATGCTTGATATTGCTGGAATTCCTCTGCCTCAG
 AACCTGAGTGGATACTCTTTGTTGCCGTTATCATCAGAAACATTTAAGAATGAACATAAAGT
 CAAAAACCTGCATCCACCCTGGATTCTGAGTGAATTCCATGGATGTAATGTGAATGCCTCCA
 CCTACATGCTTCGAACTAACCCTGGAAATATATAGCCTATTTCGGATGGTGCATCAATATTG
 CCTCAACTCTTTGATCTTTCTCGGATCCAGATGAATTAACAAATGTTGCTGTAAAATTTCC
 AGAAATTACTTATTCTTTGGATCAGAAGCTTCATTCCATTATAAACTACCCTAAAGTTTCTG
 CTTCTGTCCACCAGTATAATAAAGAGCAGTTTATCAAGTGGAAACAAAGTATAGGACAGAAT
 TATTCAAACGTTATAGCAAATCTTAGGTGGCACCAAGACTGGCAGAAGGAACCAAGGAAGTA
 TGAAAATGCAATTGATCAGTGGCTTAAACCCATATGAATCCAAGAGCAGTTT**TGA**ACAAAAA
 GTTTAAAAATAGTGTTCTAGAGATACATATAAATATATTACAAGATCATAATTATGTATTTT
 AAATGAAACAGTTTTTAATAATTACCAAGTTTTTGGCCGGGCACAGTGGCTCACACCTGTAATC
 CCAGGACTTTGGGAGGCTGAGGAAAGCAGATCACAAGGTCAAGAGATTGAGACCATCCTGGC
 CAACATGGTGAACCCCTGTCTCTACTAAAAATACAAAAATTAGCTGGGCGCGGTGGTGCACA
 CCTATAGTCTCAGCTACTCAGAGGCTGAGGCAGGAGGATCGCTTGAACCCGGGAGGCAGCAG
 TTGCAGTGAGCTGAGATTGCGCCACTGTACTCCAGCCTGGCAACAGAGTGAGACTGTGTCTGC
 AAAAAATAAAAAATAAATAATAAATTACCAATTTTTTCATTATTTTGTAAAGATGTAGTG
 TATTTTAAGATAAAATGCCAATGATTATAAAATCACATATTTTCAAAAATGGTTATTATTTA
 GGCTTTGTACAATTTCTAACAATTTAGTGGAAGTATCAAAAGGATTGAAGCAAATACTGTA
 ACAGTTATGTTCTTTAAATAATAGAGAATATAAAATATTGTAATAATATGTATCATAAAAT
 AGTTGTATGTGAGCATTTGATGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 76

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64885

<subunit 1 of 1, 536 aa, 1 stop

<MW: 61450, pI: 9.17, NX(S/T): 7

MLLLWVSVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDFSDFGRLTFHPGSQVVKLPFINE
MKTRGTSFLNAYTNSPICCPSSRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQ
KFGKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNW
LRKEAINYTEPFVIYLG LNLPHYPSPSSGENFGSSTFHTSLYWLEKVS HDAIKIPKWSPLS
EMHPVDYSSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQLDLLQKTIVIIY
SSDHGELAMEHRQFYKMSMYEASAHVPLLMMGPGIKAGLQVSNVSVSLVDIYPTMLDIAGIPL
PQNLSGYSLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYS DGAS
ILPQLFDLSSDPDEL TNVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIG
QNYSNVIANLRWHQDWQKEPRKYENAI DQWLKTHMNPRAV

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.

amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,
498-501

Sulfatases proteins:

amino acids 286-315, 359-369, 78-97

FIGURE 77

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAG
 AGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**
 GCCTCTCTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACT
 GGTGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAG
 CAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAG
 TGTGACATCTATAGCACCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGAT
 GGTGACATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCA
 CAGTCTTCTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTC
 ATCCTTGGAGGCCTCCTGGGATTCATTCTGTTGCCTGGAATCTTCATGGGATCCTACGGGA
 CTTCTACTCACCCTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGG
 GCATTATTTCTTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCCCTGCTCATCC
 CAGAGAAATCGCTCCAACCTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTC
 TCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGT
 ATGTG**TGA**AGAACCAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAG
 CACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGTGTGAGAAGGTGCTGCTGAGG
 ATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATG
 CAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTC
 CCCTGCCCTAAGTCCCCAACCCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGG
 ATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCCTG
 ACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGG
 GGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCC
 TCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTTATGACTCCACAGTGTCC
 AGACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAG
 GATGCAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

100
 90
 80
 70
 60
 50
 40
 30
 20
 10
 0

FIGURE 78

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64886
><subunit 1 of 1, 230 aa, 1 stop
><MW: 24549, pI: 8.56, NX(S/T): 1
MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGIT
QCDIYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRICTVFCQESRAKDRVAVAGGVF
FILGGLLGFIPIVAVNNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCS
SQNRNSNYDAYQAQPLATRSSPRPGQPPKVKSEFNSSYSLTGYV
```

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 79

GCACTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCC**ATG**GTGCCCAGAATCTTCGCTCCTGC
TTATGTGTCAGTCTGTCTCCTCCTCTTGTTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAG
AACCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAG
TGCTGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCCTGCAC
CTTCTGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCCCTTTGGCCTCACAAACGATTTTGTTG
TGAAGCTGAAGGTTCAAGGTGTGAATTCCCAGTGCCACTCATCTCCCATCTCCAGTAAATGT
GAAAGCAGAAGACGTTTTCCCT**TGA**AAGACATAGAAAGAAAATCAACTTTCACTAAGGCATC
TCAGAAACATAGGCTAAGGTAATATGTGTACCAGTAGAGAAGCCTGAGGAATTTACAAAATG
ATGCAGCTCCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGT
AGATTATCAGGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

FIGURE 80

MVPRI FAPAYVSVCLLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYNDAIVSLSE
TRQCGPPCTFWPCFELCCLDSFGLTNDFFVVKLVQGVNSQCHSSPISSKCESRRRFP

Signal peptide:

amino acids 1-25

1000 900 800 700 600 500 400 300 200 100 0

FIGURE 81

CTCCACTGCAACCACCCAGAGCCATGGCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATT
TTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCT
GATGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCTGCAGCACTGTT
GCTATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAACTGCACCTTCAGA
GTCTGCTTTGAGCAGTGCTGCCCCCTGGACCTTCATGGTGAAGCTGATAAACCAAGAACTGCGA
CTCAGCCCGGACCTCGGATGACAGGCTTTGTTCGAGTGTCAGCTAATGGAACATCAGGGGAA
CGATGACTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTACCTGAGA
TCTGGGATGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCACCTTCATT
CTGTGACCTGTCTGAGGCCACCCCTGCAGCTGCCCTGAGGAGGCCACAGGTCCCCTTCTAG
AATTCTGGACAGCATGAGATGCGTGTGCTGATGGGGGCCAGGGACTCTGAACCCTCCTGAT
GACCCCTATGGCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCCTCACCCT
TCTGTGAGATTTTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAA
TAAATTTATGTACTTTATAAATGAAAA

FIGURE 82

MAPRGCIVAVFAIFCISRLLC SHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVP
LARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCDSARTSDDRLCRSVS

Signal peptide:

amino acids 1-24

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 83

GGGGGCGGGTGCCTGGAGCACGGCGCTGGGGCCGCCCGCAGCGCTCACTCGCTCGCACTCAG
 TCGCGGGAGGCTTCCCCGCGCCGGCCGCGTCCCGCCCGCTCCCCGGCACCAGAAGTTCCTCT
 GCGCGTCCGACGGCGACATGGGGCGTCCCCACGGCCCTGGAGGCCGGCAGCTGGCGCTGGGGA
 TCCCTGCTCTTCGCTCTCTTCCTGGCTGCGTCCCTAGGTCCGGTGGCAGCCTTCAAGGTGCG
 CACGCCGTATTCCCTGTATGTCTGTCCCGAGGGGGCAGAACGTACCCCTCACCTGCAGGCTCT
 TGGGCCCTGTGGACAAAGGGCACGATGTGACCTTCTACAAGACGTGGTACCGCAGCTCGAGG
 GGCGAGGTGCAGACCTGCTCAGAGCGCCGGCCCATCCGCAACCTCACGTTCCAGGACCTTCA
 CCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTGGCTCAGCGCCACGGGC
 TGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATGCGCAACCTGACCCTGCTG
 GATAGCGGCCTCTACTGCTGCCTGGTGGTGGAGATCAGGCACCACCACTCGGAGCACAGGGT
 CCATGGTGCCATGGAGCTGCAGGTGCAGACAGGCCAAAGATGCACCATCCAAGTGTGTGGTGT
 ACCCATCCTCCTCCCAGGATAGTGAAAACATCACGGCTGCAGCCCTGGCTACGGGTGCCTGC
 ATCGTAGGAATCCTCTGCCTCCCCCTCATCCTGCTCCTGGTCTACAAGCAAAGGCAGGCAGC
 CTCCAACCGCCGTGCCCAGGAGCTGGTGC GGATGGACAGCAACATTCAAGGGATTGAAAACC
 CCGGCTTTGAAGCCTCACACCTGCCCAGGGGATACCCGAGGCCAAAGTCAGGCACCCCTG
 TCCTATGTGGCCCAGCGGCAGCCTTCTGAGTCTGGGCGGCATCTGCTTTCGGAGCCCAGCAC
 CCCCCTGTCTCCTCCAGGCCCCGGAGACGTCTTCTTCCCATCCCTGGACCCTGTCCCTGACT
 CTCCAAACTTTGAGGTATCTTAGCCCAGCTGGGGGACAGTGGGCTGTTGTGGCTGGGTCTGG
 GGCAGGTGCATTTGAGCCAGGGCTGGCTCTGTGAGTGGCCTCCTTGGCCTCGGCCCTGGTTC
 CCTCCCTCCTGCTCTGGGCTCAGATACTGTGACATCCCAGAAGCCCAGCCCCTCAACCCCTC
 TGGATGCTACATGGGGATGCTGGACGGCTCAGCCCCGTGTTCCAAGGATTTTGGGGTGCTGAG
 ATTCTCCCCCTAGAGACCTGAAATTCACCAGCTACAGATGCCAAATGACTTACATCTTAAGAA
 GTCTCAGAACGTCCAGCCCTTCAGCAGCTCTCGTTCTGAGACATGAGCCTTGGGATGTGGCA
 GCATCAGTGGGACAAGATGGACACTGGGCCACCCTCCCAGGCACCAGACACAGGGCACGGTG
 GAGAGACTTCTCCCCCGTGGCCGCCTTGGCTCCCCCGTTTGGCCGAGGCTGCTCTTCTGTC
 AGACTTCCTCTTTGTACCACAGTGGCTCTGGGGCCAGGCCTGCCTGCCCACTGGCCATCGCC
 ACCTTCCCCAGCTGCCTCCTACCAGCAGTTTCTCTGAAGATCTGTCAACAGGTTAAGTCAAT
 CTGGGGCTTCCACTGCCTGCATTCCAGTCCCCAGAGCTTGGTGGTCCCGAAACGGGAAGTAC
 ATATTGGGGCATGGTGGCCTCCGTGAGCAAATGGTGTCTTGGGCAATCTGAGGCCAGGACAG
 ATGTTGCCCCACCCACTGGAGATGGTGTCTGAGGGAGGTGGGTGGGGCCTTCTGGGAAGGTGA
 GTGGAGAGGGGCACCTGCCCCCGCCCTCCCCATCCCCCTACTCCCACTGCTCAGCGCGGGCC
 ATTGCAAGGGTGCCACACAATGTCTTGTCCACCCTGGGACACTTCTGAGTATGAAGCGGGAT
 GCTATTAAAACTACATGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURE 84

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64897

><subunit 1 of 1, 311 aa, 1 stop

><MW: 33908, pI: 6.87, NX(S/T): 6

MGVPTALEAGSWRWGSLFLAASLGPVAAFKVATPYSLYVCPEGQNVTLTCRLLGPVDK
GHDVTFYKTWYRSSRGEVQTCSERRPIRNLTFQDLHLHHGGHQAANTSHDLAQRHGLESASD
HHGNFSITMRNLTLTLDGLYCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNCVVYPSSSQ
DSENITAAALATGACIVGILCLPLILLVYKQQAASNRRAQELVRMDSNIQGIENPGFEAS
PPAQGIPEAKVRHPLSYVAQRQPSSEGRHLLSEPSTPLSPPGPGDVFFPSLDPVPDSPNFEVI

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 190-216

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 85

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCT
 TTCCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCC
 TTTCCTGCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGG
 GGTCTGTGGGTTGATCTGTGGCCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGA
 CTCCGCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCCTC
 TCCTCCTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCC
 AGACATGTTCTGCCTTTTCCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACT
 TGGAGCCACAAGGCCTGATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGT
 TGTACCGCCTCCACTGTCCGCCTGTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATG
 CTGTCCCAAGTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGCCCCACCAAGTCCTGCC
 AGCACAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCTCC
 CGCCTGCCCCAACCAGTGTGTCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCAC
 AACCTGCCCCGAACCAGGCTGCCCAGCACCCCTCCCCTGCCAGACTCCTGCTGCCAAGCCT
 GCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTG
 AGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGCCCGGGCACCCACAGC
 CCCCCTGCGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTTCAGACCCAAGGGAGCAG
 GCAGCACAACGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGG
 AAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCCTTGCCCTG
 CATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCACCGAGT
 ACCCCTGCCGTCACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCAGAGGACAAA
 GCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTCTCT
 CGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGG
 CCTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAG
 AGAGGTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGAATCAGATCAAGA
 AAGTCAGGAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCAC
 GAAGGTCACTGGAACGTCTTCTTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGA
 CAAAGTGACCAAGACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTT
 ATTATATATTAATAAATAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 86

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902

><subunit 1 of 1, 451 aa, 1 stop

><MW: 49675, pI: 7.15, NX(S/T): 1

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCT
CSEGAHVSCYRLHCPPVHCPQPVTEPQQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIF
SAHELFPSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEED
SVQSLHGVRHPQDPCSSDAGRKRGP GTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKH
KKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPK VAGKC
CKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKL V
KDEETEAQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPG
AEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 87

CTAGCCTGCGCCAAGGGGTAGTGAGACCGCGCGGCAACAGCTTGCGGCTGCGGGGAGCTCCC
GTGGGCGCTCCGCTGGCTGTGCAGGCGGCC**ATG**GATTCCTTGCGGAAAATGCTGATCTCAGT
CGCAATGCTGGGCGCAGGGGCTGGCGTGGGCTACGCGCTCCTCGTTATCGTGACCCCGGGAG
AGCGGCGGAAGCAGGAAATGCTAAAGGAGATGCCACTGCAGGACCCAAGGAGCAGGGAGGAG
GCGGCCAGGACCCAGCAGCTATTGCTGGCCACTCTGCAGGAGGCAGCGACCACGCAGGAGAA
CGTGGCCTGGAGGAAGAACTGGATGGTTGGCGGCGAAGGCGGCGCCAGCGGGAGGTCACCG**T**
GAGACCGGACTTGCCCTCCGTGGGCGCCGGACCTTGCGCTTGGGCGCAGGAATCCGAGGCAGCC
TTTCTCCTTCGTGGGCCCAGCGGAGAGTCCGGACCGAGATACCATGCCAGGACTCTCCGGGG
TCCTGTGAGCTGCCGTCGGGTGAGCACGTTTCCCCCAAACCCTGGACTGACTGCTTTAAGGT
CCGCAAGGCGGGCCAGGGCCGAGACGCGAGTCGGATGTGGTGAACTGAAAGAACCAATAAAA
TCATGTTCTCCAA
AAAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 92

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSS
VVHLYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQK
AIWELQVSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGQDLSTDSRTNRDMH
GLFDVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLF
FGIVGLKIFFSKFQWKIQAELDWRRKHGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFTRKSVVASQSFQAGKHYWEVDGGHNRWRVGVCRDDVDRRKEYVTLS
PDHGYWVLRRLNGEHLTYFTLNPRFISVFPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTC
RFEGLLRPYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAIPETSNSSESSSQATTPFLP
RGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

Protein Data Bank

FIGURE 93

GCGATGGTGC GCCCGGTGGCGGTGGCGGCGGCGGTTGCGGAGGCTTCCTTGGTCGGATTGCA
 ACGAGGAGAAGATGACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGC
CATGAGGAGCCTGCCGAGCCTGGGCGGCCTCGCCCTGTTGTGCTGCGCCGCCGCCGCCGCCG
 CCGTCGCCTCAGCCGCCTCGGCGGGGAATGTCACCGGTGGCGGCGGGGCCGCGGGGCAGGTG
 GACGCGTCGCCGGGCCCCGGGTTCGGGGGCGAGCCCAGCCACCCCTTCCCTAGGGGCGACGGC
 TCCCACGGCCCAGGCCCCGAGGACCGGGCCCCCGCGCGCCACCGTCCACCGACCCCTGGCTG
 CGACTTCTCCAGCCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCCTCTTCC
 ACCACCTTTTCAGGCGCCGCTCGGCCCCCTCGCCGACCACCCCTCCGGCGGCGGAACGCACTTC
 GACCACCTCTCAGGCGCCGACCAGACCCGCGCCGACCACCCCTTTCGACGACCACTGGCCCCG
 CGCCGACCACCCCTGTAGCGACCACCGTACCGGCGCCACGACTCCCCGGACCCCGACCCCC
 GATCTCCCCAGCAGCAGCAACAGCAGCGTCCTCCCCACCCACCTGCCACCGAGGCCCCCTC
 TTCGCCTCCTCCAGAGTATGTATGTAAGTGTCTGTGGTTGGAAGCCTGAATGTGAATCGCT
 GCAACCAGACCACAGGGCAGTGTGAGTGTGCGCCAGGTTATCAGGGGCTTCACTGTGAAACC
 TGCAAAGAGGGCTTTTACCTAAATTACACTTCTGGGCTCTGTCAGCCATGTGACTGTAGTCC
 ACATGGAGCTCTCAGCATACCGTGCAACAGG**TAA**GCAACAGAGGGTGGAAGTGAAGTTTATT
 TTATTTTAGCAAGGGAAAAAAAGGCTGCTACTCTCAAGGACCATACTGGTTTAAACAAAG
 GAGGATGAGGGTCATAGATTTACAAAATATTTTATATACTTTTATTCTTACTTTATATGT
 TATATTTAATGTCAGGATTTAAAAACATCTAATTTACTGATTTAGTTCTTCAAAGCACTAG
 AGTCGCCAATTTTTCTCTGGGATAATTTCTGTAAATTTTCATGGGAAAAAATTATTGAAGAAT
 AAATCTGCTTTCTGGAAGGGCTTTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTT
 ATGTTTATTAATATACCATTTGGAGTTTGAGGAAATTTGTTGTTTGGTTTATTTTTCTCTCTA
 ATCAAAATTCTACATTTGTTTCTTTGGACATCTAAAGCTTAACCTGGGGGTACCCTAATTTA
 TTTAACTAGTGGTAAGTAGACTGGTTTTACTCTATTTACCAGTACATTTTTTGAGACCAAAG
 TAGATTAAGCAGGAATTATCTTTAACTATTATGTTATTTGGAGGTAATTTAATCTAGTGGA
 ATAATGTACTGTTATCTAAGCATTTGCCTTGTACTGCACTGAAAGTAATTATTCTTTGACCT
 TATGTGAGGCACTTGGCTTTTTGTGGACCCCAAGTCAAAAACTGAAGAGACAGTATTAAAT
 AATGAAAAAATAATGACAGGTTATACTCAGTGTAACCTGGGTATAACCCAAGATCTGCTGC
 CACTTACGAGCTGTGTTCCCTTGGGCAAGTAATTTCTTTCACTGAGCTTGTTTCTTCTCAAG
 GTTGTTGTGAAGATTAATGAGTTGATATATATAAAATGCCTAGCACATGTCACTCAATAAA
 TTCTGGTTTGTTTTAAATTTCAAAGGAATATTATGGACTGAAATGAGAGAACATGTTTTAAGA
 ACTTTTAGCTCCTTGACAAAGAAGTGCTTTATACTTTAGCACTAAATATTTTAAATGCTTTA
 TAAATGATATTATACTGTTATGGAATATTGTATCATATTGTAGTTTATTAAAAATGTAGAAG
 AGGCTGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAGGCCAAGGCGGGTGGAT
 CACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAAACCCGTCTCTACTAAA
 AATACAAACAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTCGGGAGGCT
 GAGGCAGGAGAATCGGTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCTGAGATCGCGCCACT
 GCACTCCAGCCTGGTGAGAGAGGGAGACTCTGTCTTAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 94

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952
><subunit 1 of 1, 258 aa, 1 stop
><MW: 25716, pI: 8.13, NX(S/T): 5
MRS LPS LGG LALLCCAAAAA VASAASAGNVTGGGGAAGQVDASPGPGLRGEP SHPFPRATA
PTAQAPRTGPPRATVHRPLAATSPAQSPET TPLWATAGPSSTTFQAPLGPSPTTPPAAERTS
TTSQAPTRPAPTTLSTTTGPAPTTPVATTVPAPTTPRTPDLPSSSNSSVLPTPPATEAPS
SPPPEYVCNCSVVGSLNVNRCNQTTGQCECRPGYQGLHCETCKEGFYLNYSGLCQPCDCSP
HGALSIPCNR
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

N-glycosylation sites.

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

EGF-like domain cysteine pattern signature.

amino acids 214-226.

T O T E C A S E F O D I

FIGURE 96

MGGLLLAAFLALVSVPRQAVWLGRDLPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVT
LTPENNLRTLSSQHGLGGCDQSVMDLIKRN SGWVFENPSIGVLELWVLATNFRDYAIIFTQL
EFGDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

Signal peptide:

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 98

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWF
REGANTDQDAPVATNNPARAVWEETRDREHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKG
SIKWNKYKHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVS
PLDPSTTRSSVLTLLIPQPQDHGTSLTQVTFPGASVTTNKTVHLNVSYPPQNLMTVVFQGDG
TVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWV
HLRDAAEFTCRAQNPLGSQQVYLNVSLSQKATSGVTQGVVGGAGATALVFLSFCVIFVVVRS
CRKKSARPAAGVGDTGIEDANAVRGSSASQGPLEPWAEDSPPDQPPPASARSSVGEGELQYA
SLSFQMVKPWDSTRGQEATDTEYSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GACGCCCCAGTGACCTGCCGAGGTCGGCAGCACAGAGCTCTGGAG**ATGA**AAGACCCTGTTCCCTG
GGTGTACGCTCGGCCTGGCCGCTGCCCTGTCTTCACCCTGGAGGAGGAGGATATCACAGG
GACCTGGTACGTGAAGGCCATGGTGGTCGATAAAGACTTTCGGGAGGACAGGAGGCCCAGGA
AGGTGTCCCCAGTGAAGGTGACAGCCCTGGGCGGTGGGAAGTTGGAAGCCACGTTACCTTC
ATGAGGGAGGATCGGTGCATCCAGAAGAAAATCCTGATGCGGAAGACGGAGGAGCCTGGCAA
ATACAGCGCCTATGGGGGCAGGAAGCTCATGTACCTGCAGGAGCTGCCCAGGAGGGACCACT
ACATCTTTTACTGCAAAGACCAGCACCATGGGGGCCTGCTCCACATGGGAAGCTTGTGGGT
AGGAATTCTGATACCAACCGGGAGGCCCTGGAAGAATTTAAGAAATTGGTGCAGCGCAAGGG
ACTCTCGGAGGAGGACATTTTCACGCCCCTGCAGACGGGAAGCTGCGTTCCCGAACAC**TAGG**
CAGCCCCCGGGTCTGCACCTCCAGAGCCACCCCTACCACCAGACACAGAGCCCGGACCACCT
GGACCTACCCTCCAGCCATGACCCTTCCCTGCTCCCACCCACCTGACTCCAAATAAAGTCCT
TTTCCCCCAA

FIGURE 100

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65404

<subunit 1 of 1, 170 aa, 1 stop

<MW: 19457, pI: 9.10, NX(S/T): 0

MKTLFLGVTLGLAAALSFTLEEEDITGTWYVKAMVVDKDFPEDRRPRKVSPVKVTALGGGKL
EATFTFMREDRCIQKKILMRKTEEPGKYSAYGGRKLMYLQELPRRDHYIFYCKDQHHGGLLH
MGKLVGRNSDTNREALEEFKKLVQRKGLSEEDIFTPLQTGSCVPEH

Important features:**Signal peptide:**

amino acids 1-17

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 101

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCAC
AGCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCTCCAGGCCATGAGGATTCTGCAGTTAA
TCCTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAG
TGCAAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTCGAGAAGACGCGGCTACTCTGTGG
GGCGACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACA
TAGTTCACCTGGGGCAGCACAACTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCC
ACTGAGTCCTTCCCCACCCCGGCTTCAACAACAGCCTCCCCAACAAAGACCACCGCAATGA
CATCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCC
TCTCCTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCC
AGCCCCCAGTTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTGAGCACCA
GAAGTGTGAGAACGCCTACCCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGG
AAGGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCAGTCTCTT
CAAGGCATTATCTCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACAC
GAAAGTCTGCAAATATGTGGACTGGATCCAGGAGACGATGAAGAACAATTAGACTGGACCCA
CCCACCACAGCCCATCACCCCTCCATTTCCACTTGGTGTGTTGGTTCCTGTTCACTCTGTTAAT
AAGAAACCCTAAGCCAAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATG
CTGTCACTTAATAATCAACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTT
GAAATATTGTGACTCTGGGAATGACAACACCTGGTTTGTCTCTGTTGTATCCCCAGCCCCA
AAGACAGCTCCTGGCCATATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 102

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65405

<subunit 1 of 1, 250 aa, 1 stop

<MW: 27466, pI: 8.87, NX(S/T): 4

MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAH
CLKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITW
AVRPLTLSSRCVTAGTSCLISGWGSTSSPQLRLPHTLRCANITIIHQKCENAYPGNITDTM
VCASVQEGGKDSCQGDGGPLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIIQETMKNN

Important features:

Signal peptide:

amino acids 1-18

Serine proteases, trypsin family, histidine active site.

amino acids 58-63

N-glycosylation sites.

amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.

amino acids 145-148

Kringle domain proteins.

amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein

amino acids 199-209, 47-63, 220-243

Apple domain proteins

amino acids 222-249, 189-222

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 103

GAGCAGTGTTCCTGCTGGAGCCG**ATG**CCAAAAACCATGCATTTCTTATTCAGATTCATTGTTT
TCTTTTATCTGTGGGGCCTTTTTACTGCTCAGAGACAAAAGAAAGAGGAGAGCACCGAAGAA
GTGAAAATAGAAGTTTTGCATCGTCCAGAAACTGCTCTAAGACAAGCAAGAAGGGAGACCT
ACTAAATGCCCATTATGACGGCTACCTGGCTAAAGACGGCTCGAAATTCTACTGCAGCCGGA
CACAAAATGAAGGCCACCCCAAATGGTTTGTTCCTGGTGTGGGCAAGTCATAAAAGGCCTA
GACATTGCTATGACAGATATGTGCCCTGGAGAAAAGCGAAAAGTAGTTATACCCCCTTCATT
TGCATACGGAAAGGAAGGCTATGCAGAAGGCAAGATTCCACCGGATGCTACATTGATTTTTG
AGATTGAACTTTATGCTGTGACCAAAGGACCACGGAGCATTGAGACATTTAAACAAATAGAC
ATGGACAATGACAGGCAGCTCTCTAAAGCCGAGATAAACCTCTACTTGCAAAGGGAATTTGA
AAAAGATGAGAAGCCACGTGACAAGTCATATCAGGATGCAGTTTTAGAAAGATATTTTTAAGA
AGAATGACCATGATGGTGTGGCTTCATTTCTCCAAGGAATACAATGTATACCAACACGAT
GAACTA**TAG**CATATTTGTATTTCTACTTTTTTTTTTTTAGCTATTTACTGTACTTTATGTATA
AAACAAAGTCACTTTTCTCCAAGTTGTATTTGCTATTTTTCCCTATGAGAAGATATTTTGA
TCTCCCCAATACATTGATTTTGGTATAATAAATGTGAGGCTGTTTTGCAAACCTTAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 104

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65406

<subunit 1 of 1, 222 aa, 1 stop

<MW: 25794, pI: 6.24, NX(S/T): 1

MPKTMHFLFRFIVFFYLLWGLFTAQRQKKEESTEEVKIEVLHRPENCSKTSKKGDLLNAHYDG
YLAKDGSKFYCSRTQNEGHPKWFVLGVGQVIKGLDIAMTDMCPGEKRKVVIPPSFAYGKEGY
AEGKIPPDATLIFEIELYAVTKGPRSIETFKQIDMDNDRQLSKAEINLYLQREFEKDEKPRD
KSYQDAVLEDIFKKNDHDGDFISPKEYNVYQHDEL

Important features:**Endoplasmic reticulum targeting sequence.**

amino acids 219-222

N-glycosylation site.

amino acids 45-48

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 87-223, 129-142

EF-hand calcium-binding domain proteins

amino acids 202-214, 195-214

FIGURE 105

CAGAAATGCAGGGACCATTGCTTCTTCCAGGCCTCTGCTTTCTGCTGAGCCTCTTTGGAGCT
GTGACTCAGAAAACCAAAACTTCCTGTGCTAAGTGCCCCCAAATGCTTCCTGTGTCAATAA
CACTCACTGCACCTGCAACCATGGATATACTTCTGGATCTGGGCAGAACTATTCACATTCC
CCTTGGAGACATGTAACGCCAGGCATGGTGGCTCGCGCCTGTAATCCCAGTTCTTTGGGAAG
CCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATAGTGAAAC
CCCGTGTCTACTAAAAATACAAAATCAGCCGGGCGTGGTGGTGCATGCCTGCAATCCCAGT
TACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCTCAGGAGGCAGAAGTTGCAGTGAACCC
AGATCCTGCCATTGCACTCCAGCATGGATGACAGAGCAAGACTCCGTCTCAAAAAGAAAAGA
TAGTTTCTTGTTTCATTTTCGCGACTGCCCTCTCAGTGTTTCCTGGGATCCCCTCCCAAATAA
AGTACTTATATTCTC

105/249

106/249

FIGURE 106

MQGPLLLPGLCFLLSLFGAVTQKTKTSCAKCPPNASCVNNTHCTCNHGYTSGSGQKLFTFPL
ETCNARHGG SRL

Signal peptide:

amino acids 1-18

106/249

FIGURE 107

CAAGCAGGTCATCCCCTTGGTGACCTTCAAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGGCAC
AGGGAAAGGGTGACCTCTGAGATTCCCCTTTTCCCCCAGACTTTGGAAGTGACCCACC**ATGG**
GGCTCAGCATCTTTTTTGCTCCTGTGTGTTCTTGGGCTCAGCCAGGCAGCCACACCGAAGATT
TTCAATGGCACTGAGTGTGGGCGTAACTCACAGCCGTGGCAGGTGGGGCTGTTTGAGGGGCAC
CAGCCTGCGCTGCGGGGGTGTCTTATTGACCACAGGTGGGTCCTCACAGCGGCTCACTGCA
GCGGCAGCAGGTACTGGGTGCGCCTGGGGGAACACAGCCTCAGCCAGCTCGACTGGACCGAG
CAGATCCGGCACAGCGGCTTCTCTGTGACCCATCCCGGCTACCTGGGAGCCTCGACGAGCCA
CGAGCACGACCTCCGGCTGCTGCGGCTGCGCCTGCCCGTCCGCGTAACCAGCAGCGTTCAAC
CCCTGCCCCCTGCCCAATGACTGTGCAACCGCTGGCACCAGTGCCACGTCTCAGGCTGGGGC
ATCACCAACCACCCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGT
CTCCCATGCCACCTGCCATGGTGTGTATCCCGGGAGAATCACGAGCAACATGGTGTGTGCAG
GCGGCGTCCCGGGGCAGGATGCCTGCCAGGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGGA
GTCCTTCAAGGTCTGGTGTCTTGGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCCCTGG
AGTCTACACCTATATTTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAAC**TGAC**
CTGTTTCCTCCACCTCCACCCCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACC
AATATCTCCTCCATCACTTCCCCTAGCTCCACTCTTGTTGGCCTGGGAACTTCTTGGAAGTT
TAACTCCTGCCAGCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGA
ATAAATATAAATGAAGGAGGGGCAAAAAAAAAAAAAA

FIGURE 108

MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLTAAH
CSGSRYWVRLGEHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLLRLRLPVRVTSSV
QPLPLPNDCATAGTECHVSGWGITNHPRNPFDPDLLQCLNLSIVSHATCHGVYPGRITSNMVC
AGGVPGQDACQGDSGGPLVCGGVLQGLVSWGSGVGPCGQDGIPGVYTYICKYVDWIRMIMRNN

Signal peptide:

amino acids 1-17

For citation purposes only

FIGURE 109

GCGGCCACACGCAGCTAGCCGGAGCCCGGACCAGGCGCCTGTGCCTCCTCCTCGTCCCTCGC
 CGCGTCCGCGAAGCCTGGAGCCGGCGGGAGCCCCGCGCTCGCC**ATG**TCGGGCGAGCTCAGCA
 ACAGGTTCCAAGGAGGGAAGGCGTTCGGCTTGCTCAAAGCCCCGGCAGGAGAGGAGGCTGGCC
 GAGATCAACCGGGAGTTTCTGTGTGACCAGAAGTACAGTGATGAAGAGAACCTTCCAGAAAA
 GCTCACAGCCTTCAAAGAGAAGTACATGGAGTTTGACCTGAACAATGAAGGCGAGATTGACC
 TGATGTCTTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCCAAGACCCACCTGGAGATGAAG
 AAGATGATCTCAGAGGTGACAGGAGGGGTCAGTGACACTATATCCTACCGAGACTTTGTGAA
 CATGATGCTGGGGAAACGGTCGGCTGTCTCAAGTTAGTCATGATGTTTGAAGGAAAAGCCA
 ACGAGAGCAGCCCCAAGCCAGTTGGCCCCCTCCAGAGAGAGACATTGCTAGCCTGCC**TGA**
 GGACCCCGCCTGGACTCCCCAGCCTTCCCACCCCATACCTCCCTCCCGATCTTGCTGCCCTT
 CTTGACACACTGTGATCTCTCTCTCTCATTTGTTTGGTCATTGAGGGTTTGTGTTGTTT
 TCATCAATGTCTTTGTAAAGCACAAATTATCTGCCTTAAAGGGGCTCTGGGTGCGGGAATCC
 TGAGCCTTGGGTCCCCTCCCTCTCTTCTTCCCTCCTTCCCCGCTCCCTGTGCAGAAGGGCTG
 ATATCAAACCAAAAAGTAGAGGGGGCAGGGCCAGGGCAGGGAGGCTTCCAGCCTGTGTTCCC
 CTCACTTGAGGAACCAAGCACTCTCCATCCTTTCAGAAAGTCTCCAAGCCAAGTTCAGGCTC
 ACTGACCTGGCTCTGACGAGGACCCCAAGGCACTCTGAGAAGACCTTGGAGTAGGGACAAGG
 CTGCAGGGCCTCTTTCGGGTTTCCTTGGACAGTGCCATGGTTCCAGTGCTCTGGTGTCACCC
 AGGACACAGCCACTCGGGGCCCCGCTGCCCCAGCTGATCCCCACTCATTCACACCTCTTCT
 CATCCTCAGTGATGTGAAGGTGGGAAGGAAAGGAGCTTGGCATTGGGAGCCCTTCAAGAAGG
 TACCAGAAGGAACCCTCCAGTCCTGCTCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAG
 CGTGCAGCCCTACTGTCCCTTACTGGGGCAGCAGAGGGCTTCGGAGGCAGAAGTGAGGCCTG
 GGGTTTGGGGGGAAAGGTCAGCTCAGTGCTGTTCCACCTTTTAGGGAGGATACTGAGGGGAC
 CAGGATGGGAGAATGAGGAGTAAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGA
 CTGAGAAATACAAGGTTGCTTGTCTGACCCCAATCTGCTTGAAAAAAAAAAAAAAAAAAAA

MSGELSNRFQGGKAFGLLKKARQERRLAIEINREFLCDQKYSDEENLPEKLTAFKEKYMEFDLN
NEGEIDLMSLKRMMKLGVPKTHLEMKKMISEVTGGVSDTISYRDFVNMMLGKRSAVLKLV
MFEGKANESSPKPVGPPPERDIASLP

TACAGCTACAAGGGCAAGGACTGCACTGTGCCCATCAACACCTGCATCCAGAAACCCCTGTCA
GCATGGAGGCACCTGCCACCTGAGTGACAGCCACAAGGATGGGTTTCAGCTGCTCCTGCCCTC
TGGGCTTTGAGGGGCAGCGGTGTGAGATCAACCCAGATGACTGTGAGGACAACGACTGCGAA
AACAAATGCCACCTGCGTGGACGGGATCAACAACCTACGTGTGTATCTGTCCGCCTAACTACAC
AGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTCTGTCAGCATG
AGGCCAAGTGCATCCCCCTGGACAAAGGATTCAGCTGCGAGTGTGTCCCTGGCTACAGCGGC
AAGCTCTGTGAGACAGACAATGATGACTGTGTGGCCACAAGTGCCGCCACGGGGCCCAAGTG
CGTGGACACAATCAATGGCTACACATGCACCTGCCCCCAGGGCTTCAGTGGACCCTTCTGTG
AACACCCCCCACCCATGGTCCTACTGCAGACCAGCCCATGCGACCAGTACGAGTGCCAGAAC
GGGGCCCAAGTGCATCGTGGTGCAGCAGGAGCCACCTGCCGCTGCCACCAGGCTTCGCCGG
CCCCAGATGCGAGAAGCTCATCACTGTCAACTTCGTGGGCAAGACTCCTACGTGGAAGTGG
CCTCCGCCAAGGTCCGACCCCAGGCCAACATCTCCCTGCAGGTGGCCACTGACAAGGACAAC
GGCATCCTTCTCTACAAAGGAGACAATGACCCCCCTGGCACTGGAGCTGTACCAGGGCCACGT
GCGGCTGGTCTATGACAGCCTGAGTTCCTCCCAACCACAGTGTACAGTGTGAGACAGTGA
ATGATGGGCAGTTTCACAGTGTGGAGCTGGTGCAGCTAAACCAGACCCTGAACCTAGTAGTG
GACAAAGGAACTCCAAAGAGCCTGGGGAAGCTCCAGAAGCAGCCAGCAGTGGGCATCAACAG
CCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCCTCTCCGCTTGCGCCAGGGCACGG
ACCGGCCTCTAGGCGGCTTCCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAG
GACTTCAAGGCCCTCCACCCACAGTCCCTGGGGGTGTACCAGGCTGCAAGTCTGCACCGT
GTGCAAGCACGGCCTGTGCCGCTCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGCCGCCAG
GCTGGACCGGCCCACTCTGCGACCAGGAGGCCCGGGACCCCTGCCTCGGCCACAGATGCCAC
CATGGAAAATGTGTGGCAACTGGGACCTCATACATGTGCAAGTGTGCCGAGGGCTATGGAGG
GGACTTGTGTGACAACAAGAATGACTCTGCCAATGCCTGCTCAGCCTTCAAGTGTACCATG
GGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCCTGTGCCAGCCCGGCTTTAGCGGC
GAGCACTGCCAACAAGAGAATCCGTGCCCTGGGACAAGTAGTCCGAGAGGTGATCCGCGGCCA
GAAAGGTTATGCATCATGTGCCACAGCCTCCAAGGTGCCCATCATGGAATGTCTGTGGGGCT
GTGGGGCCCCAGTGCTGCCAGCCCAACCCGCAGCAAGCGGCGGAAATACGTCTTCCAGTGCACG
GACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACACTTAGAGTGCGGCTGCCTCGCGTGTTC
CTAAAGCCCCTGCCCGCCTGCCTGCCACCTCTCGGACTCCAGCTTGATGGAGTTGGGACAGCC
ATGTGGGACCCCCCTGGTGATTTCAGCATGAAGGAAATGAAGCTGGAGAGGAAGGTAAAGAAGA
AGAGAATATTAAGTATATTGTAAATAAACAAAAAATAGAACTTAAAAAATAAAAAA
AAAAAA

FIGURE 112

MAPGWAGVGAAVRARLALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRN
 AERLDLDRNNITRITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRLNKNKLQVL
 PELLFQSTPKLTRLDLSENQIQGIPRKAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEIL
 TLNNNNISRIILVTSFNHMPKIRTLRLHSNHLCDCHLAWLSDWLRQRRTVGQFTLCMAPVHL
 RGFNVADVQKKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLMEIPANLPEGIV
 EIRLEQNSIKAIPAGAFTQYKKLKRIDISKNOISDIAPDAFQGLKSLTSLVLYGNKITEIAK
 GLFDGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTISKGLFAPLQSIQTLHL
 AQNPFFVCDCHLKWLADYLQDNPIETSGARCSSPRRLANKRISQIKSKKFRCSGSEYRSRFS
 SECFMDLVCPEKCRCEGTIVDCSNQKLVRIPSHLPEYVTDLRLNDNEVSVLEATGIFKKLPN
 LRKINLSNNKIKEVREGAFDGAASVQELMLTGNQLETVHGRVFRGLSGLKTLMLRSNLISCV
 SNTDFAGLSSVRLLSLYDNRIITTITPGAFTTLVSLSTINLLSNPFNCNCHLAWLGKWLKRKR
 IVSGNPRCQKPFFLKEIPIQDVAIQDFTCDGNEESSCQLSPRCPEQCTCMETVVRCSNKGLR
 ALPRGMPKDVTELYLEGNHLTAVPRELSALRHLTLIDLSNNSISMLTNYTFSNMSHLSTLIL
 SYNRLRCIPVHAFNGLRSLRVLT LHGNDISSVPEGSFNDLTSLSHLALGTNPLHCDCSLRWL
 SEWVKAGYKEPGIARCSSPEPMADRLLLTTPTHRFQCKGPVDINIVAKCNACLSSPCKNNGT
 CTQDPVELYRCACPYSYKGKDCTVPINTCIQNPCQHGGTCHLSDSHKDGFS CSCPLGFEGQR
 CEINPDDCEDNDCENNATCVDGINNYVCICPPNYTGELCDEVIDHCVPELNLCQHEAKCIPL
 DKGFSCECVPGYSGKLCETDNDDCVAHKCRHGAQCVDTINGYTCTCPQGFSGPFCEHPPPMV
 LLQTSPCDQYECQNGAQCIVVQQEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRP
 QANISLQVATDKDNGILLYKGDNDPLALELYQGHVRLVYDSLSSPPTTVYSVETVNDGQFHS
 VELVTLNQTLNLVVDKGTPKSLGKLQKQPAVGINSPLYLGGIPTSTGLSALRQGTDRPLGGF
 HGCIEVRINNELQDFKALPPQSLGVSPGCKSCTVCKHGLCRSVEKDSVVCECRPGWTGPLC
 DQEARDPCLGHRCHHGKCVATGTSYMCKCAEGYGGDLCDNKNDSANACSAFKCHHGQCHISD
 QGEPYCLCQPGFSGEHCQQENPCLGQVVREVIRRQKGYASCATASKVPIMECRGGCGPQCCQ
 PTRSKRRKYVFQCTDGSSFVEEVERHLECGCLACS

Signal peptide:

amino acids 1-27

FIGURE 113

GGATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTTCTTCCGCA
GACTCAACTGAGAAGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCT
CCGGCAGGCTTTGAGGATGAAGGCTGCGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAG
GCGCCGAGTCCAAAATCTACACTCGTTGCAAACTGGCAAAAATATTCTCGAGGGCTGGCCTG
GACAATTACTGGGGCTTCAGCCTTGGAAGCTGGATCTGCATGGCATATTATGAGAGCGGCTA
CAACACCACAGCCCCGACGGTCCTGGATGACGGCAGCATCGACTATGGCATCTTCCAGATCA
ACAGCTTCGCGTGGTGCAGACGCGGAAAGCTGAAGGAGAACAACCACTGCCATGTCGCCTGC
TCAGCCTTGATCACTGATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGA
GACACAAGGAATGAACTATTGGCAAGGCTGGAAGAAACATTGTGAGGGCAGAGACCTGTCCG
AGTGGA AAAAAGGCTGTGAGGTTTCCTAAACTGGAAGCTGGACCCAGGATGCTTTGCAGCAAC
GCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTCATCTTGTCCCGTTTCCTCCCAATA
TTCCTTCTCAAACCTTGAGAGGGGAAAATTAAGCTATACTTTTAAGAAAATAAATATTTCCAT
TTAAATGTC

CCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTTCTTCCGCA

FIGURE 114

MKAAGILTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFSLGNWICMAYYESGYNTTAP
TVLDDGSIDYGIFQINSFAWCRRGKCLKENNHCHVACSALITDDLTDALICARKIVKETQGMN
YWQGWKKHCEGRDLSEWKKGCEVS

Signal peptide:

amino acids 1-19

115/249

FIGURE 115

CAGGCCATTTCATCCCACTGTCCTTGTGTTTCGGAGCCAGGCCACACCGTCCTCAGCAGTGT
CATGTGTTAAAAACGCCAAGCTGAATATATCCATGCCCCCTATTAAACTTGTACATGGCTCCC
CATTGGTTTTTTGGAGAAAAGTTC AAGCTTTTTACCTTGGTGTCTGCCTGTATCCCAGTGTTC
AGGCTGGCTAGACGGCGGAAGAAGATCCTATTTTACTGTCACTTCCCAGATCTGCTTCTCAC
CAAGAGAGATTCTTTTCTTAAACGACTATACAGGGCCCCAATTGACTGGATAGAGGAATACA
CCACAGGCATGGCAGACTGCATCTTAGTCAACAGCCAGTTCACAGCTGCTGTTTTTAAGGAA
ACATTCAAGTCCCTGTCTCACATAGACCCTGATGTCCTCTATCCATCTCTAAATGTCACCAG
CTTTGACTCAGTTGTTCCCTGAAAAGCTGGATGACCTAGTCCCCAAGGGGAAAAAATTCCTGC
TGCTCTCCATCAACAGATACGAAAGGAAGAAAAATCTGACTTTGGCACTGGAAGCCCTAGTA
CAGCTGCGTGGAAGATTGACATCCCAAGATTGGGAGAGGGTTCATCTGATCGTGGCAGGTGG
TTATGACGAGAGAGTCTTGAGAATGTGGAACATTATCAGGAATTGAAGAAAATGGTCCAAC
AGTCCGACCTTGGCCAGTATGTGACCTTCTTGAGGTCTTTCTCAGACAAACAGAAAATCTCC
CTCCTCCACAGCTGCACGTGTGTGCTTTACACACCAAGCAATGAGCACTTTGGCATTGTCCC
TCTGGAAGCCATGTACATGCAGTGCCAGTCATTGCTGTAAATTCGGGTGGACCCTTGGAGT
CCATTGACCACAGTGTACAGGGTTTCTGTGTGAGCCTGACCCGGTGCACCTTCTCAGAAGCA
ATAGAAAAGTTCATCCGTGAACCTTCCTTAAAGCCACCATGGGCCTGGCTGGAAGAGCCAG
AGTGAAGGAAAAAATTTTCCCCTGAAGCATTTACAGAACAGCTCTACCGATATGTTACCAAAC
TGCTGGTATTAATCAGATTGTTTTTAAGATCTCCATTAATGTCATTTTTTATGGATTGTAGACC
CAGTTTTTGAAACCAAAAAAGAAACCTAGAATCTAATGCAGAAGAGATCTTTTAAAAAATAAA
CTTGAGTCTTGAAATGTGAGCCACTTTCCTATATACCACACCTCCCTGTCCACTTTTCAGAAA
AACCATGTCTTTTATGCTATAATCATTCCAAATTTTGCCAGTGTTAAGTTACAAATGTGGTG
TCATTCCATGTTTCAGCAGAGTATTTTAATTATATTTTCTCGGGATTATTGCTCTTCTGTCTA
TAAATTTTTGAATGATACTGTGCCTTAATTGGTTTTTCATAGTTTAAGTGTGTATCATTATCAA
AGTTGATTAAATTTGGCTTCATAGTATAATGAGAGCAGGGCTATTGTAGTTCCCAGATTCAAT
CCACCGAAGTGTTCACTGTCATCTGTTAGGGAATTTTTGTTTGTCTGTCTTTGCCTGGATC
CATAGCGAGAGTGTCTGTATTTTTTTTTTAAGATAAATTTGTATTTTTTGCACACTGAGATATAA
TAAAGGTGTTTATCATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MPLLLKLVHGSPLVFGEKFKLFTLVSACIPVFRLARRRKKILFYCHFPDLLLLTKRDSFLKRLY
RAPIDWIEEYTTGMADCILVNSQFTA AVFKETFKSLSHIDPDVLYPSLNVTSFDSV VPEKLD
DLVPKGKKFLLLSINRYERKKNLTLALEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVE
HYQELKKMVQQSDLGQYVTF LRSFSDKQKISLLHSCTCVLYTPSNEHFGIVPLEAMYMQCPV
IAVNSGGPLESIDHSVTGFLCEPDPVHFSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAF
TEQLYRYVTKLLV

Signal peptide:

amino acids 1-15

117/249

FIGURE 117

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACCA**ATG**TTGGACTTCGCGATCTT
CGCCGTTACCTTCTTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGAC
AAGCTGCAGGAATTCCAGGGATTACTCCAACCTGAAGAAAAAGATGGTAATCTTCCAGATATT
GTGAATAGTGGAAGTTTGCATGAGTTCCTGGTTAATTTGCATGAGAGATATGGGCCTGTGGT
CTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTACTGAAGCAGC
ATATCAATCCCAATAAGACATCGGACCCTTTTGAAACCATGCTGAAGTCATTATTAAGGTAT
CAATCTGGTGGTGGCAGTGTGAGTGAAAACCACATGAGGAAAAAATTGTATGAAAATGGTGT
GACTGATTCTCTGAAGAGTAACTTTGCCCTCCTCCTAAAGCTTTCAGAAGAATTATTAGATA
AATGGCTCTCCTACCCAGAGACCCAGCACGTGCCCTCAGCCAGCATATGCTTGGTTTTGCT
ATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAAGATGATCAGGAAGTCATTCCG
CTTCCAGAAGAATCATGGCACAGTTTGGTCTGAGATTGGAAAAGGCTTCTAGATGGGTCAC
TTGATAAAAACATGACTCGGAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTT
TTAAGGAACATCATAAAGAACGAAAAGGAAGGAAGTTCAGTCAACATATTTTCATTGACTC
CTTAGTACAAGGGAACCTTAATGACCAACAGATCCTAGAAGACAGTATGATATTTTCTCTGG
CCAGTTGCATAATAACTGCAAAATTGTGTACCTGGGCAATCTGTTTTTTAACCACCTCTGAA
GAAGTTCAAAAAAATTATATGAAGAGATAAACCAAGTTTTTTGGAAATGGTCCTGTTACTCC
AGAGAAAATTGAGCAGCTCAGATATTGTCAGCATGTGCTTTGTGAAACTGTTTCGAACTGCCA
AACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATT
CCTAGAGAGACCCTCGTCCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCC
ATCTCCACACAAGTTTGATCCAGATCGGTTTGATGATGAATTAGTAATGAAAACTTTTTCCT
CACTTGGATTCTCAGGCACACAGGAGTGTCCAGAGTTGAGGTTTGCATATATGGTGACCACA
GTACTTCTTAGTGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAGGGACAGGTTATTGA
AACAAAGTATGAACTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGAT
AT**TAA**AATTTTATACATTTAAAATCATTGTTAAATTGATTGAGGAAAACAACCATTTAAAAA
AAATCTATGTTGAATCCTTTTATAAACAGTATCACTTTGTAATATAAACACCTATTTGTAC
TTAA

FIGURE 118

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVNSGSLHEFLVNLH
ERYGPVVSFWFGRRLVVS LGTV DVLKQHINPNKTS DPFETMLKSLLRYQSGGGSVSENHMRK
KLYENGVTDSLKSNFALLKLSEELLDKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFE
DDQEVIRFQKNHGT VWSEIGKGFLDGS LDKNMTRKKQYEDALMQLESVLRNIIKERKGRNFS
QHIFIDSLVQGNLNDQQI LEDSMIFSLASCIITAKLCTWAICFLT TSEE VQKKLYEEINQVF
GNGPVTPEKIEQLRYCQH VLCETVRTAKLTPVSAQLQDIEGKIDRFIIPRETLVL YALGVVL
QDPNTWPSPHKFD PDRFDDEL VMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLS
VEGQVIETKYELVTSSREEAWITVSKRY

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 271-290

FIGURE 119

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGA
GACCGCCGCCCTTGTCCCCGAGGGCC**ATG**GGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTC
CTGACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACAT
ACAGGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGG
TGGCCGCGCTCTCTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGA
GTCTCCATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGT
GGCCCTGTCCTTCTTCATATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTTGTCT
TCTGCAGTGCCCTTCCAGCTGTCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAA
AAGAAACCCTT**CTGA**ATTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGG
CCGCTTCGTATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTTCCCCTCGGAAACTGCTTC
TGCTGGAGGATATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATTGTATTTA
GTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACA
ATTAAAAAAAAAAAA

FIGURE 120

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLG
LFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVT
EMALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

121/249

FIGURE 121

TCCCGGACCCTGCCGCCCTGCCACT**ATG**TCCCGCCGCTCTATGCTGCTTGCCTGGGCTCTCC
CCAGCCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATA
GTGCCCCGGAACGAGTGGAAGGCCCTGGCATCAGAGTGCGCCCAGCACCTGAGCCTGCCCTT
ACGCTATGTGGTGGTATCGCACACGGCGGGCAGCAGCTGCAACACCCCCGCCTCGTGCCAGC
AGCAGGCCCCGGAATGTGCAGCACTACCACATGAAGACACTGGGCTGGTGCGACGTGGGCTAC
AACTTCCTGATTGGAGAAGACGGGCTCGTATACGAGGGCCGTGGCTGGAACCTTCACGGGTGC
CCTCAGGTCACCTTATGGAACCCCATGTCCATTGGCATCAGCTTCATGGGCAACTACATGG
ATCGGGTGCCACACCCCAGGCCATCCGGGCAGCCCAGGGTCTACTGGCCTGCGGTGTGGCT
CAGGGAGCCCTGAGGTCCAACCTATGTGCTCAAAGGACACCGGGATGTGCAGCGTACACTCTC
TCCAGGCAACCAGCTCTACCACCTCATCCAGAATTGGCCACACTACCGCTCCCCC**TG**AGGCC
CTGCTGATCCGCACCCCATTCTCCCTCCCATGGCCAAAACCCCACTGTCTCCTTCTCCA
ATAAAGATGTAGCTC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 122

MSRRSMLLAWALPSLLRLGAAQETEDPACCSPIVPRNEWKALASECAQHLSLPLRYVVVSHT
AGSSCNTPASCQQQARNVQHYHMKTLGWCDVGYNFLIGEDGLVYEGRGWNFTGAHSGHLWNP
MSIGISFMGNYMDRVPTPQAIRAAQGLLACGVAQGALRSNYVLKGHRDVQRTLSPGNQLYHL
IQNWPHYRSP

Signal peptide:

amino acids 1-20

1043304

FIGURE 123

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAG**ATG**CAACT
 GACTCGCTGCTGCTTCGTGTTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGG
 ATGATGGTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCC
 CGGGTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCT
 CCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCCTTGGGCAGCCCCCAACCGCC
 CGAACCACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTAC
 TCCAACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAA
 TGGGACCTTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCC
 TCGTGCCCCCAGTAAAGCTGTAGAGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAG
 GCCTCCAAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCCGCCGGACCTC
 GCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCAGACCACGCTCAGAGCTCAGCCACCT
 GGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTAT
 CGGCTGGTCCAGAAGGTGTGCCCAGATTACAACCTACCATAGTGATACCCCCTACTACCCATC
 TGGG**TGA**CCCCGGGGCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGC
 AGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTCAGGCAGGGAGGGGGGTGGAG
 ACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCCAAGTG
 CTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGGAGGAGGAGTGGGCT
 CTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCC
 TGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGGTT
 CTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACAGCCTG
 TCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGGC
 CAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATG
 GCATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTC
 CAGCCAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATT
 GCTGATGGCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCCACACTAAGGCC
 ACAGCCCATCCGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAG
 CATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCC
 CGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGA
 GCATGTGCTGGATCTGTTCTGTGTGTCTGTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGT
 GAAACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTT
 GCCCCGGGGCA

FIGURE 124

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521

><subunit 1 of 1, 252 aa, 1 stop

><MW: 28127, pI: 8.91, NX(S/T): 5

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMAN
STLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVD
HNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGR
RTSLCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPY
YPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

Downloaded from www.cambridge.org/core

FIGURE 125

GTGAATGTGAGGGTTTGATGACTTTTCAGATGTCTAGGAACCAGAGTGGGTGCAGGGGCCCA
GGCAGGGCTGATTCTTGGGCGGAGGAGAGTAGGGTAAAGGGTTCTGCATGAGCTCCTTAAAG
GACAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGACTTTGACTTCAAGCCACAGAAT
TGGTGGAAGTGTGCGCGCCGCCGCCGCTCGCTCCTGCAGCGCTGTGACCTAGCCGCTAG
CATCTTCCCAGACACGGGATCCCGGGGTAGGAGGCGACGCGGGCGAGCACCAGCGCCAGCC
GGCTGCGGCTGCCCACACGGCTCACCATGGGCTCCGGGCGCCGGGCGCTGTCCGCGGTGCCG
GCCGTGCTGCTGGTCCTCACGCTGCCGGGGCTGCCCGTCTGGGCACAGAACGACACGGAGCC
CATCGTGCTGGAGGGCAAGTGTCTGGTGGTGTGCGACTCGAACCCGGCCACGGACTCCAAGG
GCTCCTCTTCCCTCCCCGCTGGGGATATCGGTCCGGGCGGCCAACTCCAAGGTCGCCTTCTCG
GCGGTGCGGAGCACCAACCACGAGCCATCCGAGATGAGCAACAAGACGCGCATCATTTACTT
CGATCAGATCCTGGTGAATGTGGGTAATTTTTTTCACATTGGAGTCTGTCTTTGTAGCACCAA
GAAAAGGAATTTACAGTTTTCAGTTTTCACGTGATTAAAGTCTACCAGAGCCAACTATCCAG
GTTAACTTGATGTTAAATGGAAAACCAAGTAATATCTGCCTTTGCGGGGGACAAAGATGTTAC
TCGTGAAGCTGCCACGAATGGTGTCTGCTCTACCTAGATAAAGAGGATAAGGTTTACCTAA
AACTGGAGAAAGGTAATTTGGTTGGAGGCTGGCAGTATTCACGTTTTCTGGCTTTCTGGTG
TTCCCCCTATAGGATTCAATTTCTCCATGATGTTTCATCCAGGTGAGGGATGACCCACTCCTG
AGTTATTGGAAGATCATTTTTTTCATCATTTGGATTGATGTCTTTTATTGGTTTCTCATGGGTG
GATATGGATTCTAAGGATTCTAGCCTGTCTGAACCAATACAAAATTTACAGATTATTTGTG
TGTGTCTGTTTCAGTATATTTGGATTGGGACTCTAAGCAGATAAATACCTATGCTTAAATGTA
ACAGTCAAAAGCTGTCTGCAAGACTTATTCTGAATTTCAATTCCTGGGATTACTGAATTAGT
TACAGATGTGGAATTTTATTTGTTTAGTTTTAAAGACTGGCAACCAGGTCTAAGGATTAGA
AACTCTAAAGTTCTGACTTCAATCAACGGTTAGTGTGATACTGCCAAAGAACTGTATACTG
TGTTAATATATTGATTATATTTGTTTTTATTCCCTTTGGAATTAGTTTGTTTGGTTCTTGTA
AAAACCTGGATTTTTTTTTTTCAGTAACTGGTATTATGTTTTCTCTTAAAATAAGGTAATGAA
TGGCTTGCCCACAAATTTACCTTGACTACGATATCATCGACATGACTTCTCTCAAAAAAAAA
GAATGCTTCATAGTTGTATTTTAATTGTATATGTGAAAGAGTCATATTTTCCAAGTTATATT
TTCTAAGAAGAAGAATAGATCATAAATCTGACAAGGAAAAAGTTGCTTACCCAAAATCTAAG
TGCTCAATCCCTGAGCCTCAGCAAAACAGCTCCCCCTCCGAGGGAAATCTTATACTTTATTGC
TCAACTTTAATTAAATGATTGATAATAACCACTTTATTAAAAACCTAAGGTTTTTTTTTTTT
TCCGTAGACATGACCACCTTTATTAACCTGGTGGTGGGATGCTGTTGTTTCTAATTATACCTAT
TTTTCAAGGCTTCTGTTGTATTTGAAGTATCATCTGGTTTTGCTTAACTCTTTAAATTGTA
TATATTTATCTGTTTAGCTAATATTAAATTCAAATATCCCATATCTAAATTTAGTGCAATAT
CTTGTCTTTTGTATAGGTCATATGAATTCATAAAATTATTTATGTCTGTTATAGAATAAAGA
TTAATATATGTTAAAAAAA

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTC
GCCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGC
TTTCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTA
TTGACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTC
TATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTT
GAAGAGTATAAACCCAGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCT
TGGGCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGG
CCAGGCACAGTGGGCATTCATGGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGAC
GCTGGTCATTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGA
AAAAGTGGGGCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTGAGCCAGACCTTC
ATAAGTTCTTATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCATGGGCAC
CTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAG
ACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGATTAACCTCAGGGAACCAGCACTTCCCAA
ACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCTTTTTTCTGAAAATCCCTTTTTCTG
GTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 128

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658
><subunit 1 of 1, 257 aa, 1 stop
><MW: 28472, pI: 9.33, NX(S/T): 0
MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDN
KDGPTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGF
GIMSGVFSFVNTLSDSLGPSTVGIHGDSPQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKKW
GILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKN
FLLYNQSR
```

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 129

CGGCAACCAGCCGCCACCACCGCTGCCACTGCCGCCCTGCCGGGGCCATGTTTCGCTCTGGGCTTGCCCTTCT
 TGGTGCTCTTGGTGGCCTCGGTGAGAGCCATCTGGGGGTTCTGGGGCCCAAGAACGTCTCGCAGAAAGACGCCG
 AGTTTGAGCGCACCTACGTGGACGAGGTCAACAGCGAGCTGGTCAACATCTACACCTTCAACCATACTGTGACCC
 GCAACAGGACAGAGGGCGTGCCTGTGTCTGTGAACGTCTGAACAAGCAGAAGGGGGCGCCGTGCTGTTTGTGG
 TCCGCCAGAAGGAGGCTGTGGTGTCTTCCAGGTGCCCTAATCCTGCGAGGGATGTTTCAGCGCAAGTACCTCT
 ACCAAAAAGTGGAACGAACCCTGTGTGACGCCCCCACCAGAATGAGTCGGAGATTGAGTTCTTCTACGTGGATG
 TGTCCACCCTGTCAACAGTCAACACCACATACCAGCTCCGGGTGACCCGATGGACGATTTTGTGCTCAGGACTG
 GGGAGCAGTTTCACTTCAATACCACAGCAGCAGCCCCAGTACTTCAAGTATGAGTTCCCTGAAGGCGTGGACT
 CGGTAATTGTCAAGGTGACCTCCAACAAGGCCCTTCCCTTGCTCAGTCATCTCCATTGAGGATGTGCTGTGTCTG
 TCTATGACCTGGACAACAACGTAGCCTTCATCGGCATGTACCAGACGATGACCAAGAAGGGCGGCCATCACCGTAC
 AGCGCAAAGACTTCCCCAGCAACAGCTTTTATGTGGTGGTGGTGGTGAAGACCGAAGACCAAGCCTGCGGGGGCT
 CCCTGCCTTTCTACCCCTTTCGAGAAAGATGAACCGGTGATCAAGGGCACC GCCAGAAAACCCTGTGAGTGCTGG
 TGTCTCAAGCAGTCACTCTGAGGCATACGTGAGTGGGATGCTCTTTTGCTGGGTATATTTCTCTCCTTTTACC
 TGCTGACCGTCCCTGCGCTGCTGGGAGAATGGAGGCAGAAGAAGACCCCTGCTGGTGGCCATTGACCGAG
 CCTGCCCAGAAAGCGGTCAACCTCGAGTCTGGCTGATTCTTTTCTGGCAGTTCCCTTATGAGGGTTACAAC
 ATGGCTCCTTTGAGAATGTTTCTGGATCTACCGATGGTCTGGTTGACAGCGCTGGCACTGGGGACCTCTCTTACG
 GTTACCAGGGCGCTCCTTTGAACCTGTAGTACTCGCCCCGAGTGGACTCCATGAGCTCTGTGGAGGAGGATG
 ACTACGACACATTGACCGACATCGATTCCGACAAGAATGTCATTTCGACCAAGCAATACCTCTATGTGGCTGACC
 TGGCAGGGAAGGACAAGCGTGTCTGCGGAAAAAGTACCAGATCTACTTCTGGAACATTGCCACCATTGCTGTCT
 TCTATGCCCTTCCCTGCTGGTGCAGCTGGTATCACCACCAGCGGTGGTGAATGTACAGGGAATCAGGACATCT
 GCTACTACAACCTTCTCTGCGCCCAACCACTGGGCAATCTCAGCGCCTTCAACAACATCCTCAGCAACCTGGGGT
 ACATCCTGCTGGGGCTGCTTTTCTGCTCATCATCCTGCAACGGGAGATCAACCACAACCGGGCCCTGCTGCGCA
 ATGACCTCTGTGCCCTGGAATGTGGGATCCCCAAACACTTTGGGCTTTTCTACGCCATGGGCACAGCCCTGATGA
 TGGAGGGGGCTGCTCAGTGCTTGTATCATGTGTGCCCCAACTATAACCAATTTCCAGTTTGACACATCGTTTATGT
 ACATGATCGCCGGACTCTGCATGCTGAAGCTCTACCAGAAGCGGCACCCGGACATCAACGCCAGCGCCTACAGTG
 CCTACGCCCTGCCCTGGCCATTGTCATCTTCTCTGTGCTGGGCGTGGTCTTTGGCAAAGGGGAACACGGCGTTCT
 GGATCGTCTTCTCCATCATTCACATCATCGCCACCCCTGCTCCTCAGCAGCAGCTCTATTACATGGGCCGGTGG
 AACTGGACTCGGGGATCTTCCGCCGCATCCTCCACGTGCTCTACACAGACTGCATCCGGCAGTGACGCGGGCCGC
 TCTACGTGGACCGCATGGTGTGCTGGTCTGTCATGGGCAACGTGTCATCAACTGGTGGTGGCTGCCTATGGGCTTATCA
 TGGCCCCCAATGATTTTCGCTTCTTCTTCTTCCAGGGACTCAGCAGTCCGAGAAAACCCCTGCAGAGTCGAGGGAGACA
 ACCGGGACTGCATCCTCCTCGACTTCTTTGACGACCACGACATCTGGCACTTCTCTCCTCCTATCGCCATGTTTCG
 GGTCTTCTGGTGTGCTGACACTGGATGACGACCTGGATACTGTGACGCGGGACAAGATCTATGTCTTCTAGC
 AGGAGCTGGGGCCCTTCGCTTCACTCAAGGGGCCCTGAGCTCCTTTGTGTATAGACCGGTCACTCTGTCTGTGCT
 GTGGGGATGAGTCCCAGCACCCTGCCAGCACTGGATGGCAGCAGGACAGCCAGGTCTAGCTTAGGCTTGGCCT
 GGGACAGCCATGGGGTGGCATGGAACCTTGCAGCTGCCCTCTGCCGAGGAGCAGGCCTGCTCCCCTGGAACCCCC
 AGATGTTGGCCAAATTGCTGCTTTCTTCTCAGTGTTGGGGCCTTCCATGGGGCCCTGTCTTTGGCTCTCCATTT
 GTCCCTTTGCAAGAGGAAGGATGGAAGGGACACCCTCCCCATTTTATGCCTTGCATTTTGGCCGTCTCCTCCCC
 ACAATGCCCCAGCCTGGGACCTAAGGCCTCTTTTTCTCCATACTCCCACTCCAGGGCCTAGTCTGGGGCCTGA
 ATCTCTGTCTGTATCAGGGCCCCAGTTCTCTTTGGGCTGTCCCTGGCTGCCATCACTGCCATTCCAGTCAGCC
 AGGATGGATGGGGGTATGAGATTTTGGGGGTGGCCAGCTGGTGCCAGACTTTTGGTGCTAAGGCCTGCAAGGGG
 CCTGGGGCAGTGCGTATTTCTTCCCTCTGACCTGTGCTCAGGGCTGGCTCTTTAGCAATGCGCTCAGCCCAATT
 TGAGAACCCTTCTGATTCAAGAGGCTGAATTCAGAGGTCACTCTTTCATCCCATCAGCTCCAGACTGATGCC
 AGCACCAGGACTGGAGGGAGAAGCGCTCACCCCTTCCCTTCTTCTTCCAGGCCCTTAGTCTTGCCAAACCCC
 AGCTGGTGGCCTTTTCACTGCCATTGACACTGCCCAAGAATGTCCAGGGGCAAGGAGGGATGATACAGAGTTTCA
 CCCGTTCTGCTCCACAGCTGTGGGCACCCCACTGCTTACCTTAGAAAGGGGCTTCAGGAAGGGATGTGCTGTTT
 CCCTCTACGTGCCAGTCTTAGCCTCGCTCTAGGACCCAGGGCTGGCTTCTAAGTTTCCGTCCAGTCTTCCAGGCA
 AGTTCTGTGTAGTCATGCACACATACCTATGAACCTTGGAGTTTACAAAGAATTGCCCCAGCTCTGGGCAC
 CCTGGCCACCCTGGTCTTGGATCCCCCTTGGTCCACCTGGTCCACCCAGATGCTGAGGATGGGGGAGCTCAGG
 CGGGGCTCTGCTTTGGGGATGGGAATGTGTTTTTCTCCAAACTTGTTTTTATAGCTCTGCTTGAAGGGCTGGG
 AGATGAGGTGGTCTGGATCTTTCTCAGAGCGTCTCCATGCTATGGTTGCATTTCCGTTTTCTATGAATGAATT
 TGCATTCAATAAACAACCAGACTCAAAAAAAAAAAAAA

FIGURE 130

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66659
><subunit 1 of 1, 832 aa, 1 stop
><MW: 94454, pI: 6.94, NX(S/T): 12
MFALGLPFLVLLVASVESHLGVLGPKNVSQKDAEFERTYVDEVNSELVNIYTFNHTVTRNRT
EGVRVSVNVLNKQKGAPLLFVVRQKEAVVSFQVPLILRGMFQRKYLYQKVERTLCQPPTKNE
SEIQFFYVDVSTLSPVNTTYQLRVSRMDDFVLRTGEQFSFNTTAAQPQYFKYEFPEGVDSVI
VKVTSNKAFFPCSVISIQDVLCPVYDLNNAVFIGMYQMTKKAAITVQRKDFPSNSFYVVVV
VKTEDQACGGSPLPFYFPAEDEPVDQGHRQKTLVSVLSQAVTSEAYVSGMLFCLGIFLSFYLL
TVLLACWENWRQKKKTLLVAIDRACPESGHPRVLADSFPGSSPYEGYNYGSFENVSGSTDGL
VDSAGTGDLSTYGYQGRSFEPVGTRPRVDSMSSVEEDDYDTLTDIDSDKNVIRTKQYLYVADL
ARKDKRVLRRKKYQIYFWNIATIAVFYALPVVQLVITYQTVVNVNTGNQDICYYNFLCAHPLGN
LSAFNNILSNLGYILLGLLFLLIILQREINHNRALLRNDLCALECGIPKHFGFLFYAMGTALM
MEGLLSACYHVCPTYNTNFQFDTSEFMYMIAGLCMLKLYQKRHPDINASAYSAYACLAIVIFFS
VLGVVFGKGNTAFWIVFSIIHIIATLLSTQLLYMGRWKLDSGIFRRILHVLYTDCIRQCSG
PLYVDRMVLLVMGNVINWSLAAYGLIMRPNDFASYLLAIGICNLLLYFAFYIIMKLRSGERI
KLIPLLCIVCTSVVWGFALFFFFQGLSTWQKTPAESREHNRDCILLDFDHDHIWHFLSSIA
MFGSFLVLLTLDDDLDTVQRDKIYVF
```

Important features of the protein:**Signal peptide:**

amino acids 1-18

Transmembrane domains:

amino acids 292-317, 451-470, 501-520, 607-627, 751-770

Leucine zipper pattern.

amino acids 497-518

N-glycosylation sites.

amino acids 27-30, 54-57, 60-63, 123-126, 141-144, 165-168, 364-367, 476-479, 496-499, 572-575, 603-606, 699-702

FIGURE 131

GCTCAAGTATGCCCTGCCCTTGCCCCACCCAGCCAGCCAGCTGGCCAGAGAGGCCCTGGAGAAAGGAGC
 TCTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTCTG
 ACC**ATG**GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGC
 CCAGCCTGCAGAGCTGTCTGTGGAAGTTCCAGAAAACCTATGGTGGAAATTTCCCTTTATACC
 TGACCAAGTTGCCGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCA
 GGCAAGGCAACTGAGGGGCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAG
 GGCCCTGGACCGAGAGGAGCAGGCAGAGTACCAGCTACAGGTCACCCTGGAGATGCAGGATG
 GACATGTCTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTC
 CCCCATTTCTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCC
 CTTTCTTCTTCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTTCGAT
 TCCACATCCTGAGCCAGGCTCCAGCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGG
 CTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGAC
 CTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTG
 CCACCGTGGAAGTCTCCATCATAGAGACACCTGGGTGTCCCTAGAGCCTATCCACCTGGCA
 GAGAATCTCAAAGTCTTATACCCGCACCATGCCCCAGGTACACTGGAGTGGGGGTGATGT
 GCACTATCACCTGGAGAGCCATCCCCCGGACCCTTTGAAGTGAATGCAGAGGGAAACCTCT
 ACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAG
 AATTCCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTGATGGATGAGAA
 TGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCAC
 CAGGTAAGTGAAGTGAAGTACTGTGTCAGCAGAGGATGCAGATGCCCCCGGCTCCCCCAATTCC
 CACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTCCA
 GGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAAC
 TCCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGT
 GAAGTGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGAT
 TGGGCCCTATAAGCCTCCCTGAGGATGTGGAGCCCGGAGCACTGCTGGTGGCCATGCTAACAGCCA
 TTGATGCTGACCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACA
 GAAGGGACTTTTGGCCTGGATTGGGAGCCAGACTCTGGGCATGTGTAGACTCAGACTCTGCAA
 GAACCTCAGTTATGAGGCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGA
 AGCTGGTGGGGCCAGGCCAGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGA
 GTGATGCCACCCCCCAAGTTGGACCAGGAGAGCTACGAGGCCAGTGTCCCATCAGTGCCCC
 AGCCGGCTCTTTCTGCTGACCATCCAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCT
 CCTAGTCAATGACTCAGAGGGCTGGCTCTGCATTGAGAAATTCTCCGGGGAGGTGCACACC
 GCCCAGTCCCTGCAGGGCGCCAGCCTGGGGACACCTACACGGTGCTTGTGGAGGCCCAGGA
 TACAGCCCTGACTCTTGCCCCTGTGCCCTCCCAATACCTCTGCACACCCCGCCAAGACCATG
 GCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCTGGCCAGTGGGCACGGTCCCTACAGC
 TTCACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGCGCCTCCAGACTCTCAATGGTTCC
 CCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGTGAACACATAATCCCCGTGG
 TGGTCAGCCACAATGCCCAGATGTGGCAGCTCCTGGTTGAGTGTGCTGTGCTGCAAC
 GTGGAGGGGAGTGCATGCGCAAGGTGGGCCGCATGAAGGGCATGCCACGAAGCTGTGCGC
 AGTGGGCATCCTTGTAGGCACCTGGTAGCAATAGGAATCTTCTCATCCTCATTTTACACC
 ACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCCTGAAGGCG
 ACTGTCT**TGA**ATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAGTCCC
 CTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCAT
 CTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAAC
 TTTATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTGTTGCCCAATAATAAAGCCCCA
 GAGAACTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 132

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSG
KATEGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVP
HFSQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSDFQLEPRL
GALALSPKGSTSLDHALERTYQLLVQVKMDQASGHQATATVEVSIIESTWVSLEPIHLAE
NLKVLYPHHMAQVHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQN
SHGEDYAAPLELHVLVMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSH
VVYQLLSPEPEDGVEGRAFQVDPTSGSVTLGVLPRLAGQNILLVLAMDLAGAEGGFSSTCE
VEVAVTDINDHAPEFITSQIGPISLPEDVEPGTLVAMLTADLEPAFRLMDFAIERGDTE
GTFGLDWEPPDSGHVRLRLCKNLSYEAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERV
MPPPKLDQESYEASVPISAPAGSFLLTIQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTA
QSLQGAQPGDITYTVLVEAQDTALTAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSF
TLGPNPTVQRDWRLQTLNGSHAYLTALHWWVEPREHIIPVVVSHNAQMWQLLVRVIVCRCNV
EGQCMRKVGRMKGMPTKLSAVGILVGTLLVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 133

CCGGGGGAC**ATG**AGGTGGATACTGTTCAATTGGGGCCCTTATTGGGTCAGCATCTGTGGCCAA
GAAAAATTTTTTGGGGACCAAGTTTTGAGGATTAATGTCAGAAATGGAGACGAGATCAGCAA
ATTGAGTCAACTAGTGAATTCAAACAACCTTGAAGCTCAATTTCTGGAAATCTCCCTCCTCCT
TCAATCGGCCTGTGGATGTCCTGGTCCCATCTGTCAGTCTGCAGGCATTTAAATCCTTCCTG
AGATCCCAGGGCTTAGAGTACGCAGTGACAATTGAGGACCTGCAGGCCCTTTTAGACAATGA
AGATGATGAAATGCAACACAATGAAGGGCAAGAACGGAGCAGTAATAACTTCAACTACGGGG
CTTACCATTCCCTGGAAGCTATTTACCACGAGATGGACAACATTGCCGCAGACTTTCCTGAC
CTGGCGAGGAGGGTGAAGATTGGACATTCGTTTGAAAACCGGCCGATGTATGTACTGAAGT
CAGCACTGGGAAAGGCGTGAGGCGGCCGGCCGTTTGGCTGAATGCAGGCATCCATTCCCGAG
AGTGGATCTCCCAGGCCACTGCAATCTGGACGGCAAGGAAGATTGTATCTGATTACCAGAGG
EATCCAGCTATCACCTCCATCTTGGAGAAAATGGATATTTTTCTTGTTGCCTGTGGCCAATCC
TGATGGATATGTGTATACTCAAACCTCAAACCGATTATGGAGGAAGACGCGGTCCCGAAATC
CTGGAAGCTCCTGCATTGGTGCTGACCCAAATAGAAACTGGAACGCTAGTTTTTGAGGAAAG
GGAGCCAGCGACAACCCCTTGCTCCGAAGTGTACCATGGACCCACGCCAATTCGGAAGTGG
GGTGAAATCAGTGGTAGATTTTCATCCAAAAACATGGGAATTTCAAGGGCTTCATCGACCTGC
ACAGCTACTCGCAGCTGCTGATGTATCCATATGGGTACTCAGTCAAAAAGGCCCCAGATGCC
GAGGAACTCGACAAGGTGGCGAGGCTTGCGGCCAAAGCTCTGGCTTCTGTGTGGGGCACTGA
GTACCAAGTGGGTCCCACCTGCACCACTGTCTATCCAGCTAGCGGGAGCAGCATCGACTGGG
CGTATGACAACGGCATCAAATTTGCATTACATTTGAGTTGAGAGATACCGGGACCTATGGC
TTCCTCCTGCCAGCTAACCAGATCATCCCCACTGCAGAGGAGACGTGGCTGGGGCTGAAGAC
CATCATGGAGCATGTGCGGGACAACCTCTAC**TAG**GCGATGGCTCTGCTCTGTCTACATTTAT
TTGTACCCACACGTGCACGCACTGAGGCCATTGTTAAAGGAGCTCTTTCCTACCTGTGTGAG
TCAGAGCCCTCTGGGTTTGTGGAGCACACAGGCCTGCCCTCTCCAGCCAGCTCCCTGGAGT
CGTGTGTCCTGGCGGTGTCCTGCAAGAACTGGTTCTGCCAGCCTGCTCAATTTTGGTCCTG
CTGTTTTTGTATGAGCCTTTTGTCTGTTTCTCCTTCCACCCTGCTGGCTGGGCGGCTGCACTC
AGCATACCCCTTCTGGGTGGCATGTCTCTCTCTACCTCATTTTTTAGAACCAAAGAACATC
TGAGATGATTCTCTACCTCATCCACATCTAGCCAAGCCAGTGACCTTGCTCTGGTGGCACT
GTGGGAGACACCACCTTGCTTTTAGGTGGGTCTCAAAGATGATGTAGAATTTCTTTAATTTT
TCGCAGTCTTCTGGAAAATATTTTTCTTTGAGCAGCAAATCTTGTAGGGATATCAGTGAAG
GTCTCTCCCTCCCTCCTCTCCTGTTTTTTTTTTTTTTTTTGGAGACAGAGTTTTGCTCTTGTTGCC
CAGGCTGGAGTGTGATGGCTCGATCTTGGCTCACCACAACCTCTGCCTCCTGGGTTCAGGA
ATTCTCCTGCCTCAGCCTCTTGAGTAGCTTGGTTTATAGGCGCATGCCACCATGCCTGGCTA
ATTTTTGTGTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGTCAAGGCTGGTCTCAAACCTCCCA
ACCTCAGGTGATCTGCCCTCCTTGGCCTCCCAGAGTGCTGGGATTACAGGTGTGAGCCACTG
TGCCGGGGCCCGTCCCCTCCTTTTTTTAGGCCTGAATACAAAGTAGAAGATCACTTTCTTTCAC
TGTGCTGAGAATTTCTAGATACTACAGTTCTTACTCCTCTCTTCCCTTTGTTATTCACTGTG
ACCAGGATGGCGGGAGGGGATCTGTGTCACTGTAGGTACTGTGCCCAGGAAGGCTGGGTGAA
GTGACCATCTAAATTGCAGGATGGTGAAATTATCCCCATCTGTCTTAATGGGCTTACCTCCT
CTTTGCCTTTTTGAACCTCACTTCAAAGATCTAGGCCTCATCTTACAGGTCTTAAATCACTCAT
CTGGCCTGGATAATCTCACTGCCCTGGCACATTCCCATTTGTGCTGTGGTGTATCCTGTGTT
TCCTTGTCTGGTTTTGT
TCTGTCTATTTTGTATCCTGGACCACAAGTTCCTAAGTAGAGCAAGAATTCATCAACCAGCT
GCCTCTTGTTCATTTACCTCAGCACGTACCATCTGTCTTTTGTGTTGTTGTTGTTTGTGTT
TTGTTTTTTTTGCTTTTACCAAACATGTCTGTAAATCTTAACCTCCTGCCTAGGATTTGTACA
GCATCTGGTGTGTGCTTATAAGCCAATAAATATTCAATGTGAAAAAAAAAAAAAAAAAAAA

FIGURE 134

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKLNFWKSPSSFNR
PVDVLVPSVSLQAFKSFLRSQGLEAYVTIEDLQALLDNEDDEMQHNEGQERSNNFNYGAYH
SLEAIYHEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRRPAVWLNAGIHSREWI
SQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANPDGYVYTQTQNRLLWRKTRSRNPGS
SCIGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKSVVDFIQKHGNFKGFIDLHSY
SQLLMYPYGYSVKKAPDAEELDKVARLAALKALASVSGTEYQVGPTCTTVYPASGSSIDWAYD
NGIKFAFTFELRDTGTYGFLLPANQIIPTAEETWLGLKTIMEHVRDONLY

Signal peptide:

amino acids 1-16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 135

[illegible]

FIGURE 136

MASYLYGVLFVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLVLE
TPSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVP
SKDLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQ GK
VVDIIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAF
GVDTELNCFVLQMDYKGDVAFFVLPSKGKMRQLEQALSARTLIKWSHSLQKRWIEVFIPRF
SISASYNLETILPKMGIQNAFDKNADFSGIAKRDSLQVSKATHKAVLDVSEEGTEATAATTT
KFIVRSKDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

Signal peptide:

amino acids 1-20

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

FIGURE 138

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATI
SGSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSE
SSTPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSS
TTSSGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTT
SNGAGTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSS
GASTATNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGA
NTATNSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEAST
ATNSESSTVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAV
SEAKPGGSLVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGP
GPGGNHGAPHRPRWSPNWFWRPVS SIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 139

GGGAGAGAGGATAAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTC
CCAGCAAT**ATG**CATCTTGACGCTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCC
CTGTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAACCGAGG
GCTGAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGAATCACGC
ATGCCGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGC
AAGGAGTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGAT
CAACCATGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACG
CTGCTGGACAGGCCGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCAC
CAGGCTGGGAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGG
AAAGGAAGTGGAGAAGCTTGGCCAAGGTGCCACCATGCTGCTGGCCAGGCCGGAAGGAGC
TGCAGAATGCTCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGC
AACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGG
GGCCTCAGTCAACACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCA
TGCCCT**TAA**ACTGGCATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGAC
ATGACCTGGAGGGGTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTTGTACTG
GGATTTGTGAATAAACTTGATACACCA

p.f.c.f. e.e.s.f.d.f.

FIGURE 140

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675
><subunit 1 of 1, 247 aa, 1 stop
><MW: 25335, pI: 7.00, NX(S/T): 0
MHLARLVGSCSLLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAG
REVEKVFNGLSNMGSHGTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAG
QAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQN
AHNGVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

141/249

FIGURE 141

CTCCGGGTCCCCAGGGGCTGCGCCGGGCGGCCTGGCAAGGGGGACGAGTCAGTGGACACTCCAGGAAGAGCGGC
 CCCGCGGGGGCGATGACCGTGCGCTGACCCTGACTCACTCCAGGTCCGGAGGCGGGGGCCCCGGGGCGACTCG
 GGGGCGGACCGCGGGGCGGAGCTGCCGCCCGTGAGTCCGGCCGAGCCACCTGAGCCCCGAGCCGCGGGACACCGTC
 GCTCCTGCTCTCCGAATGCTGCGCACCGCGATGGGCCTGAGGAGCTGGCTCGCCGCCCATGGGGCGGCTGCCG
 CCTCGGCCACCGCTGCTGCTGCTCCTGCTGCTGCTGCTCCTGCTGCAAGCGCCGCTCCGACCTGGGCGCTCAGC
 CCCCCGATCAGCCTGCCTCTGGGCTCTGAAGAGCGGCCATTCTCAGATTGCAAGCTGAACACATCTCCAACCTAC
 ACAGCCCTTCTGCTGAGCAGGATGGCAGGACCCTGTACGTGGGTGCTCGAGAGGCCCTCTTTGCACTCAGTAGC
 AACCTCAGCTTCTGCCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGACAGCGAGAGAAGAAACAGCAGTGC
 AGCTTCAAGGGCAAGGACCCACAGCGGACTGTCAAACTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCAC
 CTGTTTACCTGTGGCACAGCAGCCTTTCAGCCCCATGTGTACCTACATCAACATGGAGAATTTCACCCTGGCAAGG
 GACGAGAAGGGGAATGTCTCCTGGAAGATGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCC
 CTGGTGGTTGATGGCGAGCTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGACCCGGCCATCTCGCGGAGC
 CAAAGCCTTCGCCCCACCAAGACCGAGAGCTCCCTCAACTGGCTGCAAGACCCAGCTTTTGTGGCCTCAGCCTAC
 ATTCTGAGAGCCTGGGCAGCTTGCAAGGCGATGATGACAAGATCTACTTTTTCTTCAGCGAGACTGGCCAGGAA
 TTTGAGTTCTTTGAGAACCATTGTGTCCCGCATTTGCCCGCATCTGCAAGGGCGATGAGGGTGGAGAGCGGGTG
 CTACAGCAGCGCTGGACCTCCTTCTCAAGGCCAGCTGCTGTGCTCAGGCCCCGACGATGGCTTCCCTTCAAC
 GTGCTGCAGGATGTCTTACGCTGAGCCCCAGCCCCCAGGACTGGCGTGACACCCTTTTCTATGGGGTCTTCACT
 TCCCAGTGGCACAGGGGAATACAGAAGGCTCTGCCGTCTGTGTCTTCACAATGAAGGATGTGCAGAGAGTCTTC
 AGCGGCCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCGGTGCCACACCCCGG
 CCTGGAGCGTGCATCACCAACAGTGCCCGGGAAAGGAAGATCAACTATCCCTGCAGCTCCCAGACCGCGTGCT
 AACTTCTCAAGGACCACTTCTGATGGACGGGCGAGTCCGAAGCCGATGCTGCTGCTGCAGCCCCAGGCTCGC
 TACCAGCGCTGGCTGTACACCGCTCCCTGGCCTGCACCACACCTACGATGTCTCTTCTTGGGCACTGGTGAC
 GGCCGGCTCCACAAGGCAGTGAGCGTGGGCCCCCGGGTGACATCATTGAGGAGCTGCAGATCTTCTCATCGGGA
 CAGCCCGTGCAAGATCTGCTCCTGGACACCCACAGGGGGCTGCTGTATGCGGCCTCACACTCGGGCGTAGTCCAG
 GTGCCATGGCCAACTGCAGCCTGTACCGGAGCTGTGGGGAGTGCCTCCTCGCCCCGGGACCCCTACTGTGCTTGG
 AGCGGCTCCAGCTGCAAGCACCTCAGCCTTACCAGCCTCAGCTGACCCAGGCGGTGGATCCAGGACATCGAG
 GGAGCCAGCGCCAAGGACCTTTGCAAGCGCTCTTCGGTTGTGTCCCCGTCTTTTGTACCAACAGGGGAGAAGCCA
 TGTGAGCAAGTCCAGTTCAGCCCAACACAGTGAACACTTTGGCCTGCCCGCTCCTCTCCAACCTGGCGACCCGA
 CTCTGGCTACGCAACGGGGCCCCCGTCAATGCCTCGGCCTCCTGCCACGTGCTACCCACTGGGGACCTGCTGCTG
 GTGGGCACCCAACAGCTGGGGGAGTTCCAGTGTGCTGCTCACTAGAGGAGGGCTTCCAGCAGCTGGTAGCCAGCTAC
 TGCCAGAGGTGGTGGAGACGGGGTGGCAGACCAACAGATGAGGGTGGCAGTGTACCCGTCAATTATCAGCACA
 TCGCGTGTGATGTCACCAAGCTGGTGGCAAGGCCAGCTGGGGTGACAGAGTCTACTGGAAGGAGTTCCTGGTG
 ATGTGCACGCTCTTTGTGCTGGCCGTGCTGCTCCAGTTTTATTCTTGTCTTACCAGGACCCGGAACAGCATGAAA
 GTCTTCTGAAGCAGGGGAATGTGCCAGCGTGACCCCCAAGACCTGCCCTGTGGTGCTGCCCCCTGAGACCCGC
 CCACTCAACGGCTAGGGCCCCCTAGCACCCCGCTCGATCACCGAGGGTACCAGTCCCTGTGAGACAGCCCCCG
 GGGGCCCCGAGTCTTCACTGAGTCAGAGAAGAGGGCACTCAGCATCCAAGACAGCTTCGTGGAGGTATCCCCAGTG
 TGCCCCCGCCCCGGGTCCGCCTTGGCTCGGAGATCCGTGACTCTGTGGTGTGAGAGCTGACTTCCAGAGGACGC
 TGCCCTGGCTTCAGGGGCTGTGAATGCTCGGAGAGGGTCAACTGGACCTCCCCTCCGCTCTGCTCTTCGTGGAAC
 ACGACCGTGGTGGCCGGCCCTTGGGAGCCTTGGAGCCAGCTGGCCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCC
 TACCACCCAGACACCAACAGCCGTGGCCCCAGAGGTCTTGCCAAATATGGGGGCTGCCTAGGTGGTGGAA
 CAGTGCTCCTTATGTAACTGAGCCCTTTGTTTAAAAACAATTCCAATGTGAACTAGAATGAGAGGGAAGAG
 ATAGCATGGCATGCAGCACACAGGCTGCTCCAGTTCATGGCCTCCCAGGGGTGCTGGGGATGCATCCAAAGTGG
 TTGTCTGAGACAGAGTTGGAACCCCTCACCACTGGCCTCTTACCTTCCACATTATCCCGCTGCCACCGGCTGC
 CCTGTCTCACTGCAGATTGAGGACAGCTTGGGCTGCGTGCGTTCTGCCTTGCCAGTCAGCCGAGGATGTAGTTG
 TTGCTGCCGTGCTCCACACCTCAGGGACCAGAGGGCTAGGTTGGCACTGCGGCCCTCACCAGGTCTTGGGCTC
 GGACCCAACCTCTGGACCTTTCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACAGCGCGAGCTCAGGAGAGA
 TTTCTGTGACAATGTACGCCTTTCCCTCAGAATTCAGGGAAGAGACTGTGCGCTGCCTTCTCCGTTGTGCGTGA
 GAACCCGTGTGCCCTTCCACCATATCCACCTCGCTCCATCTTTGAACTCAAACACGAGGAACCTAACTGCACC
 CTGGTCTCTCCCCAGTCCCCAGTTTACCCTCATCCCTCACCTTCTCCACTCTAAGGGATATCAACACTGCCC
 AGCACAGGGGCCCTGAATTTATGTGTTTTTATACATTTTTTAATAAGATGCACCTTATGTCAATTTTTTAATAAA
 GTCTGAAGAATTACTGTTTTAAAAA

FIGURE 142

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962

><subunit 1 of 1, 837 aa, 1 stop

><MW: 92750, pI: 7.04, NX(S/T): 6

MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLLQPPPPTWALSPRISLPLGSEERPFLRF
EAEHISNYTALLLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEKKQQCSFKG
KDPQRDCQNYIKILLPLSGSHLFTCGTAAFSPMCTYINMENFTLARDEKGNVLLEDGKGRCP
FDPNFKSTALVVDGELYTGTVSSSQGNPAISRSQSLRPTKTESSLNWLQDPAFVASAYIPE
SLGSLQGDDDKIYFFFSETGQEFFEFFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLC
SRPDDGFPFENVLQDVFTLSPPQDWRDTLFYGVFTSQWHRGTTEGSAVCVFTMKDVQRVFSG
LYKEVNRETQQWYTVTHPVPTPRPGACITNSARERKINSSQLPDRVLNFLKDHFLMDGQVR
SRMLLLQPQARYQRVAVHRVPGLHHTYDVLFLGTGDGRLHKAVSVGPRVHIIEELQIFSSGQ
PVQNLLLDTHRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSQYQP
QLATRPWIQDIEGASAKDLCSASSVSPSFVPTGEKPCQVQFQPNVTNLTACPLLSNLATR
LWLRNGAPVNASASCHVLPTGDLLLVTGQQLGEFQCWSLEEGFQQLVASYCPEVVEDGVADQ
TDEGGSVPVPIISTSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSM
KVFLKQGECAVHPKTCPVVLPPETRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTESEKR
PLSIQDSFVEVSPVCPRPRVRLGSEIRDSVV

Transmembrane domains:

amino acids 23-46 (type II), 718-738

CTAAGCCGAGGAGTGTGCAGCTGCGGCGCGCGCGCGCGGCTACGAAGAGGACGGGACAGGCGCCGTGCGAACCGA
GCCAGCCAGCCGAGGACGCGGGCAGGGCGGGACGGGAGCCCGGACTCGTCTGCCGCGCGCGTCTGCGCGTCTG
TGCCGGCCCCCGCTCCCCGCGCGGAGCGGGAGGAGCCGCCACCTCGCGCCCCGAGCCGCCGCTAGCGCGCGC
CGGGCATGGTCCCCTCTTAAAGGCGCAGCGCCGCGCGCGCGCGGGCGGGGTGTGCGGAACAAAGCGCCGGCGCGGGG
CCTGCGGGCGGCTCGGGGCGCGCATGGCGCGCGCGCGCGCGCGCGCGCGCGCTGCCCGGGCGGGGCTCG
CGGCGCTAGGGCGGGCTGGCTCCGTGGCGGGGGGACGCGGCTGAGGGCGCGCGGAGCTCGGGCGCGCGGGC
GGCGCGCGCGCGCGCCCCGCGGGCGGAGCGCGCGGGC**ATG**CGCGCGCGCGCGCGCGCTGGCTCAGCGTGC
TGCTCGGGCTCGTCTGGGCTTCTGTGCTGGCTCGCGGCTCGTCTGCCCGGGCTTCCGAGCTGAAGCGAGCGG
GCCACCGCGCGCGCGCAGCCCCGAGGGCTGCCGCTCCGGGCAGGCGGCGGCTTCCAGGGCCGGCGGGGCGCGCG
CGCATGCGCGCGGGCGCAGCTCTGGCCGCCCGGCTCGGACCCAGATGGCGCGCCCCGCGCGACAGGAACCTTCTCT
TCGTGGGAGTCAAGCCGCCAGAAATACCTCGAGACTCGGGCGGTGGCCGCTACAGAACTGGTCCAAGACAA
TTCCTGGGAAAGTTTCAAGTCTTCTCAAGTAGGGTTCTGACACATCTGTACCAATCCAGTAGTGCCACTACGGG
GTGTGGACGACTCCTACCCGCCCCAGAAGAAGTCTTTCATGATGCTCAAGTACATGCACGACCCTACTTGCACA
AGTATGAATGGTTTATGAGAGCAGATGATGACGTGTACATCAAAGGAGACCGTCTGGAGAACTTCTGAGGAGTT
TGAACAGCAGCGAGCGCCCTCTTTCTTGGGCGAGACAGGCCTGGGCACCACGGAAGAAATGGGAAAACCTGGCCCTGG
AGCCTGGTGAGAACTTCTGCATGGGGGGCGCTGGCGTGATCATGAGCCGGGAGGTGCTTCGGAGAATGGTGGCCG
ACATTGGCAAGTGTCTCCGGGAGATGTACACCACCTGAGGACGTGGAGGTGGGAAGGTGTGTCGGGAGGTTT
CAGGGGTGCAGTGTGTCTGGTCTTATGAGATGCGGCAGCTTTTTTATGAGAATTACGAGTGAACAACAAAAGGGT
ACATTAGAGATCTCCATAACAGTAAATTCACCAAGCTATCACATTACACCCCAACAAAAACCCACCTACCAGT
ACAGGCTCCACAGCTACATGCTGAGCCGCAAGATATCCGAGCTCCGCCATCGCACAAATACAGCTGCACCGCGAAA
TTGTCCTGATGAGCAAAATACAGCAACACAGAAATTCATAAAGAGGACCTCCAGCTGGGAATCCCTCCCTCCTCA
TGAGGTTTACGCCCCCAGCAGAGGAGATCTGGAATGGGAGTTTCTGACTGGAAAATACTTGTATTTCGGCAG
TTGACGGCCAGCCCCCTCGAAGAGGAATGGACTCGGCCACGAGGGAAGCTTGGACGACATGTCTCAGCGTCA
TGGAGATGATCAATGCCAACGCCAAGACCAGAGGGCGCATCATTGACTTCAAAGAGATCCAGTACGGCTACGCC
GGGTGAACCCCATGTATGGGGCTGAGTACATCCTGGACCTGCTGCTTCTGTACAAAAAGCACAAAGGGAAGAAAA
TGACGGTCCCTGTGAGGAGGCACGCGTATTTACAGCAGACTTTTACGAAAATCCAGTTTGTGGAGCATGAGGAGC
TGGATGCACAAGAGTTGGCCAAGAGAATCAATCAGGAATCTGGATCCTTGCTCCTTCTCTCAAACCTCCCTGAAGA
AGCTCGTCCCCTTTACGCTCCCTGGTCTGCAAGAGTGAGCACAAGAAACCAAGATAAAAGATAAACATACTGA
TTCTTTTGTCTGGGCGTTTCGACATGTTTGTGAGATTTATGGGAACTTTGAGAAGACGTGTCTATCCCCAATC
AGAAGCTCAAGCTCGTGGTTCTGCTTTTCAATTCTGACTCCAACCTTGACAAGGCCAAACAAGTTGAACTGATGA
GAGATTACCGCATTAAGTACCCTAAAGCCGACATGCAGATTTTGCCTGTGTCTGGAGAGTTTCAAGAGCCCTGG
CCCTGGAAGTAGGATCCTCCAGTTTAACAATGAATCTTGTCTTCTTCTGCGACGTCGACCTCGTGTTTACTA
CAGAATTCCTTCAGCATCTCGACAAAATACAGTTCTGGGCAACAAATATTTTCCAATCATCTTCAGCCAGT
ATGACCCAAAGATTGTTTATAGTGGGAAAGTTCCCAGTGACAACCAATTTTGCCTTTACTCAGAAAACCTGGCTCT
GGAGAACTATGGGTTTGGCATCACGTGTATTTATAAGGGAGATCTTGTCCGAGTGGGTGGCTTTGATGTTTCCA
TCCAAGGCTGGGGGCTGGAGGATGTGGACCTTTTCAACAAGTTGTCCAGGCAGGTTTGAAGACGTTTAGGAGCC
AGGAAGTAGGAGTAGTCCACGTCCACCATCCTGTCTTTTGATGCCAATCTTGACCCCAACAGTACAAAATGT
GCTTGGGTGCAAAAGCATCGACCTATGGTCCACCCAGCAGCTGGCTGAGATGTGGCTGGAAAAAATGATCCAA
GTTACAGTAAAGCAGCAATAATAATGGCTCAGTGAGGACAGCC**TAA**TGTCCAGCTTGTGCTGGAAAGACGTTTT
TAATTATCTAATTTATTTTTCAAAAAATTTTTTGTATGATCAGTTTTTGAAGTCCGTTATACAAGGATATATTTTAC
AAGTGGTTTTCTTACATAGGACTCCTTTAAGATTGAGCTTTCTGAACAAGAAGGTGATCAGTGTGTGCTTTGAA
CACATCTTCTTGCTGAACATTATGTAGCAGACCTGCTTAACCTTTGACTTGAATGTACCTGATGAACAAAACCTT
TTTAAAAAATGTTTTCTTTTGGAGACCTTTTGCTCCAGTCTTATGGCAGAAAACGTGAACATTCCTGCAAAAGTAT
TATTTAACAACAAACACTGTAACCTGGTAATGTTCTGTTGTATTGTTAACTTCCACAGATTCTACCTTTTGT
GTTTTGTTTTTTTTTTTTTACAATTGTTTTAAAGCCATTCTGATGTTCCAGTTGTAAGATAAGGAAATGTGATAATA
GCTGTTTCATCATTGTCTTCAGGAGAGCTTTCAGAGTTGATCATTTCTCTCATGGTACTCTGCTCAGCATGGC
CAGTAGGTTTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGAGACGGAGTCTCACTCTGTTACCCAGGCTGGAATG
CAGTGGCGCAATCTTGGCTCACTTTTAACTCCAATCTCCCTGGTTCAAGCAATTTCCCTGCCTTTGCCTCCCGAGT
AGCTGGGATTACAGGCACACACCCACGCCAGNTAGTTTTTTTTGTATTTTGTAGTAGAGCGGGGTTTACCAT
GCAAGCCCAGCTGGCCACGTAGGTTTTTAAAGCAAGGGGCGTGAAGAAGGCACAGTGAGGTATGTGGCTGTTCTCG
TGGTAGTTTATTCGGCCTAAATAGACCTGGCATTAAATTTCAAGAAGGATTTGGCATTCTCTCTTTCAGCCCT
CTCTTTAAAGGGTAAAATATTAATGTTTAGAATGACAAAGATGAATTATTACAATAAATCTGATGTACACAGACT
GAAACATACACACATACACCTTAATCAAAACGTTGGGGAAAAATGTATTTGGTTTTTGTTCCTTTTCATCCTGTCTG
TGTATGTGGGTGGAGATGGTTTTTCACTTCTTCACTACTGTTTTTCTTTATCTTCTTGTATCTGAAATACCTTTAA
TTTTATTTAATATCTGTTGTTTCAGAGCTCTGCCATTTCTTGAGTACCTGTTAGTTAGTATTATTTATGTGTATCGG
GAGTGTGTTTAGTCTGTTTTATTTGCAGTAAACCGATCTCCAAAGATTTCTTTTTGGAAACGCTTTTTCCCTCC

TTAATTTTATATTCTTACTGTTTTACTAAATATTAAGTGTTCTTTGACAATTTTGGTGCTCATGTGTTTTGGG
GACAAAAGTGAAATGAATCTGTCAATTATACCAGAAAGTTAAATTCTCAGATCAAATGTGCCTTAATAAATTTGTT
TTCATTTAGATTTCAAACAGTGATAGACTTGCCATTTTAATACACGTCAATGGAGGGCTGCCGTAATTTGTAAATAG
CCTGATGCTCATTGGAAAAATAAACCAAGTGAACAATTTTTCTATTGTACTTTTTGCAACCAATTTGTCTCATT
ATTCCTGTTTTAGCTGAAGAATTGTATTACATTTGGAGAGTAAAAAAGTTAAACACGAAAAAA

FIGURE 144

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836

><subunit 1 of 1, 802 aa, 1 stop

><MW: 91812, pI: 9.52, NX(S/T): 3

MAARGRRRAWLSVLLGLVLGFVLASRLVLPRASELK RAGPRRRASPEGCRSGQAAASQAGGAR
GDARGAQLWPPGSDPDGGPRDRNFLFVGVM T AQKYLQTRAVAA YRTWSKTIPGKVQFFSSEG
SDTSVPIPVVPLRGVDDSYPPQKKSFMMLKYMHDHYLDKYEFWMRADD DVYIKGDRL ENFLR
SLNSSEPLFLGQTGLGTTEEMGKLAL EPGENFCMGGPGVIMSREVLRRMVP HIGKCLREMYT
THEDVEVGRCVRRFAGVQCVWSYEMRQLFYENYEQNKKGYIRDLHNSKI HQAITLHPNKNPP
YQYRLHSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMR FQPRQREE
ILEWEFLTGKYLYSAVDGQPPRRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQY
GYRRVNPMYGA EYILDLLLLYKKHKGKKMTVPVRRHAYLQQTFSKIQFVEHEELDAQELAKR
INQESGSLSF LSNLKKLVPFQLPGSKSEHKEPKDKKINILIPLSGRFDMFVRFMGNFEKTC
LIPNQNVKLVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQ
FNNESLLFFCDVDLVFTTEFLQRCRANTVLGQQIYFPIIFSQYDPKIVYSGKVPSDNHFAFT
QKTGFWRNYGFGITCIYKGD LVRVGGFDVSIQGWGLEDDVDFNKVVQAGLKTFRSQEVGVVH
VHHPVFCDPNLDPKQYKMCLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA

Signal peptide:

amino acids 1-23

FIGURE 145

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCC
TTTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCT
ACGGACGACGCCT**ATGA**AGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTG
TGCCGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAACTTGAATCATTATATA
CAAGTTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAAATC
TAACTCTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAAATTTAAGGAGCTAGTTA
CACATGGAGACGCTTCAACTGAGAATGATGTTTTAACCAATCCTATCAGTGAAGAACTACA
ACTTTCCCTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAAACACACGGAAAGTACCCC
ATTCTGGTCGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTG
AAAATGAAGAGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAACTGAGGCACCAAGAATG
TTGCCAGTTGTTACTGAATCATCTACAAGTCCATATGTTACCTCATACAAGTCACCTGTCAC
CACTTTAGATAAGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCCTCAGCTCT
CAGGTGAAACTGCGATAGAAAAACCCGAAGAGTTTGGAAGCACCCAGAGAGTTGGAATAAT
GATGACATTTTGAAAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACTTCTTAGTGA
CACCAGCAACCCAGCATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCC
TTGCTCTAGCAGCAGCAGCAGAACATAAATTAAAAACAATGTATAAGTCCCAGTTATTGCCA
GTAGGACGAACAAGTAATAAAATTGATGACATCGAACTGTTATTAACATGCTGTGTGAATTC
TAGATCTAAACTCTATGAATATTTAGATATTAAATGTGTTCCACCAGAGATGAGAGAAAAAG
CTGCTACAGTATTCAATACATTAAAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTA
AAAGTTTAT**TAA**ACAATAATATAAAAAATTTTAAACCTACTTGATATTCCATAACAAAGCTGA
TTTAAGCAAACCTGCATTTTTTTCACAGGAGAAATAATCATATTCGTAATTTCAAAGTTGTAT
AAAAATATTTTCTATTGTAGTTCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACA
ATTTTCATATGCACTAAAAACCTAATTTAAATAAAATTTTGGTTCAGGAAAAA

FIGURE 146

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68864

><subunit 1 of 1, 350 aa, 1 stop

><MW: 39003, pI: 5.59, NX(S/T): 1

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRSPSGEPGREKKSNSPK
HVYSIASKGSKFKELVTHGDASTENDVLTNPISEETTTFTPTGGFTPEIGKKKHTESTPFWSI
KPNNVSIVLHAEOPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSSTSPYVTSYKSPVTTLTK
STGIEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNP
AYREDIEASKDHLKRSLALAAAAEHKLKTMYSQLLPVGRTSNKIDDIETVINMLCNSRSKL
YEYLDIKCVPPPEMREKAATVFNTLKNMCRSRRVTALLKVY

Signal peptide:

amino acids 1-19

148/249

FIGURE 147

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTG
ACCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACA
GGAATATCC**ATG**GCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACA
GTGGCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCT
CCTGCTCCCTCTTTCCCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAG
TTCCATGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACA
GTATCGAGGGGAGAACTGAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGC
TAAAAACATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTCAGTTCCCAGATTTAC
GATGAGGAGGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCAT
CGTGGGATATGTTGACGGAGGTATCCAGTTACTCTGCCTGTCCTCAGGCTGGTTCCCCCAGC
CCACAGCCAAGTGGAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCA
GATGGGTACAGCCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCAT
ATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAG
AGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGT
GCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAATCCA
GGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACG
CAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAA
ACTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAA
GAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGAC
AAAATGTAGGGTGGTATGTGGGAGTGTGTGCGGGATGACGTAGACAGGGGGGAAGAACAATGTG
ACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTTAC
ATTCAATCCCCATTTTATCAGCCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCC
TGGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACC
CTGCTGACATGTCAGTTTGAAGGCTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGA
GGAAAAGGGGACTCCCATATTCATATGTCCAGTGTCTTGGGGAT**TG**AGACAGAGAAGACCCTG
CTTAAAGGGCCCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGC
CCCAGCTTCCTCTCCGGAGCCTGCGCACAGAGAGTCACGCCCCCCTCTCCTTTAGGGAGC
TGAGGTTCTTCTGCCCTGAGCCCTGCAGCAGCGGCAGTCACAGCTTCAGATGAGGGGGGAT
TGGCCTGACCCTGTGGGAGTCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCA
CATTAGGTTTAGTTTGTGAAAACCTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCC
CAGGCTCCTCATTTGCTAGTCACGGACAGTGATTCTTGCCTCACAGGTGAAGATTAAAGAGA
CAACGAATGTGAATCATGCTTGCAGTTTGGAGGGCACAGTGTTTGCTAATGATGTGTTTTTA
TATTATACATTTTCCACCATAAACTCTGTTTGCTTATTCCACATTAATTTACTTTTCTCTA
TACCAAATCACCCATGGAATAGTTATTGAACACCTGCTTTGTGAGGCTCAAAGAATAAAGAG
GAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCATTACCTGATACCAAACAGGCAAAG
AAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCCCTCATTAACACAGACACAAAAA
TTCTAAATAAAATTTTAAACAAATTAACTAAACAATATATTTAAAGATGATATATAACTACT
CAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAAATATTTAAATATCAACCAGTGTAATT
CAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAA

FIGURE 148

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866

><subunit 1 of 1, 466 aa, 1 stop

><MW: 52279, pI: 6.16, NX(S/T): 2

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHA
VVHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEE
ATWELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGY
SLYDVEISIIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALC
GVVMGMIIVFFKSKGKIQAELDWRRKHGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFTTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDGRGKNNVTLS
PNNGYWVLRRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLT
CQFEGLLRPYIQHAMYDEEKGTPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 149

CCTTCACAGGACTCTTCATTGCTGGTTGGCA**ATG**ATGTATCGGCCAGATGTGGTGAGGGCTA
 GGAAAAGAGTTTGTGGGAACCCCTGGGTATCGGCCTCGTCATCTTCATATCCCTGATTGTC
 CTGGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAAGAAGACCTACAA
 TTACTATAGCACATTGTCATTTACAACCTGACAACTATATGCTGAGTTTGGCAGAGAGGCTT
 CTAACAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAA
 TCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTATCAAGTTCAGTCAACAGAAGCATGG
 AGTGTTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAG
 ATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTA
 GATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCA
 TTGCTGCGGAACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAG
 AAGTAGAAGAGGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGT
 GGAGCAACCTTAATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAA
 GAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACAATAAAACCTTCGAAAATGAAACGGG
 GTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCT
 CTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGA
 TGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAA
 ATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAAC
 TGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTT
 AGAAGGAAAAACAGATGCATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTA
 GAGATATCTGGTACCTTGCTGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCAACAAG
 CCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTATC**TA**
AGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATT
 TTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAAC
 TGTTTGCTTGATGCATGTATTTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTG
 CCAGATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATA
 ATACAATATTACATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTGTCAGAATTTTGAC
 TTGTTGACATAAAATTTGTAATGCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTC
 AGCTCCTCTCATTTTCAGCAAATATCCATTTTCAAGGTGCAGAACAAAGGAGTGAAAGAAAATA
 TAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGT
 GGAATATTAGAAATGATCATATTATTATGAAAGGTCAAGCAAAGACAGCAGAATACCAATC
 ACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAGATATA
 TCCTTATTTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTG
 ACCTATAATAATTATACAACTTCATGCAATGTACTTGTTCTAAGCAAATTAAGCAAATAT
 TTATTTAACATTGTTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 150

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871
><subunit 1 of 1, 423 aa, 1 stop
><MW: 47696, pI: 8.96, NX(S/T): 3
MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTD
KLYAEFGREASNNFTEMSQRLESMVKNAFYKSPREEFVKSQVIKFSQQKHGVLAHMLLICR
FHSTEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTL
GQSLRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFG
VTIKPSKMKRGLRRIIVHEKYKHPSHDYDISLAEISSPVPYTNAVHRVCLPDASYEFQPGDV
MFVTGFGALKNDGYSQNHRLQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGD
SGGPLVSSDARDIYWLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI
```

Transmembrane domain:

amino acids 21-40 (type II)

GTCGAAGGTTATAAAAGCTTCCAGGCCAAACGGCATTGAAGTTGAAGATACAACCTGACAGCA
 CAGCCTGAGATCTTGGGGATCCCTCAGCCTAACACCCACAGACGTCAGCTGGTGGATTCCCC
 CTGCATCAAGGCCTACCCACTGTCTCCATGCTGGGCTCTCCCTGCCTTCTGTGGCTCCTGGC
 CGTGACCTTCTTGGTTCCAGAGCTCAGCCCTTGGCCCCCTCAAGACTTTGAAGAAGAGGAGG
 CAGATGAGACTGAGACGGCGTGGCCGCCTTTGCCGGCTGTCCCCTGCGACTACGACCACTGC
 CGACACCTGCAGGTGCCCTGCAAGGAGCTACAGAGGGTCGGGCCGGCGGCCTGCCTGTGCCC
 AGGACTCTCCAGCCCCGCCAGCCGCCGACCCGCCGCGCATGGGAGAAGTGCGCATTTGCGG
 CCGAAGAGGGCCGCGCAGTGGTCCACTGGTGTGCCCCCTTCTCCCCGGTCTTCCACTACTGG
 CTGCTGCTTTGGGACGGCAGCGAGGCTGCGCAGAAGGGGCCCCCGCTGAACGCTACGGTCCG
 CAGAGCCGAAGTGAAGGGGCTGAAGCCAGGGGGCATTTATGTGCTTTGCGTAGTGGCCGCTA
 ACGAGGGCCGGGGCAAGCCGCGTGCCCCAGGGCTGGAGGAGAGGGCCTCGAGGGGGCCGACATC
 CCTGCCTTCGGGCCTTGACAGCCGCCTTGCGGTGCCGCCCAACCCCCGCACTCTGGTCCACGC
 GGCCGTGCGGGTGGGCACGGCCCTGGCCCTGCTAAGCTGTGCCGCCCTGGTGTGGCACTTCT
 GCCTGCGCGATCGCTGGGGCTGCCCGCGCCGAGCCGCCGCCCGAGCCGCAGGGGCGCTCTGA
 AAGGGGCCTGGGGGCATCTCGGGCACAGACAGCCCCACCTGGGGCGCTCAGCCTGGCCCCCG
 GGAAAGAGGAAAACCCGCTGCCTCCAGGGAGGGCTGGACGGCGAGCTGGGAGCCAGCCCCAG
 GCTCCAGGGCCACGGCGGAGTCATGGTTCTCAGGACTGAGCGCTTGTTTAGGTCCGGTACTT
 GGCGCTTTGTTTCCTGGCTGAGGTCTGGGAAGGAATAGAAAGGGGGCCCCCAATTTTTTTTTTA
 AGCGGCCAGATAATAAATAATGTAACCTTTGCGGTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 152

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68874
><subunit 1 of 1, 238 aa, 1 stop
><MW: 25262, pI: 6.44, NX(S/T): 1
MLGSPCLLWLLAVTFLVPRAQPLAPQDFEEEEADETETAWPPLPAVPCDYDHCRHLQVPCKE
LQRVGPAAACLC PGLSSPAQPPDPPRMGEVRIAAEEGRAVVHWCAFPSPVLHYWLLLWDGSEA
AQKGPPLNATVRR AELKGLKPGGIYVVCVVAANEAGASRVPQAGGEGLEGADIPAFGPCSRL
AVPPNPRTL VHAAVGVGTALALLSCAALVWHFCLRDWGCPRRAAARAAGAL
```

Important features of the protein:**Signal peptide:**

amino acids 1-20

Transmembrane domain:

amino acids 194-220

N-glycosylation site.

amino acids 132-135

FIGURE 154

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880
><subunit 1 of 1, 305 aa, 1 stop
><MW: 35383, pI: 5.99, NX(S/T): 0
MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPV
VHPVMIAVCCFLIIVGMLGYCGTVKRNLLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQ
WSDMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREF
PGCSKQAHQEDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWAL
YYDRREPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL
```

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 155

GAGAGAGGCGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCC
TGCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGA
CCTGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAG
ACCGGGAGGATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGC
CTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGAT
CCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCCTCATCAAGG
TGATTCTGGATAAAATACTACTTCCCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAG
CTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTT
CCCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGG
ACTCGGCCACAGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAG
ACAGCCTGTAGGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGA
TGTTGTTGAAATCACAGAAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTC
TCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTGTGGGAAGAGCCTGAAGACCCCCCGT
GTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTCAGCATCCAGTACGA
CAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCCTCACGGCAGCCCACT
GCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGC
AGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAA
AGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGC
CCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGA
TGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCA
GGTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAAGA
TGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCC
CTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGG
GGGCCCCGAGCACCCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATG
TCTGGAAGGCTGAGCTG**TAA**TGCTGCTGCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTG
CCCTGCCCCACCTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACCC
CTCTGCCCCACAGCCTCAGCATTTCTTGGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCC
TCGCAGCCCAGAGGCGCCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCC
AGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAA
GGAACCTTTCACACACTACTGAATGGAAGCAGGCTGTCTTGTAAGAGCCAGATCACTGTGGG
CTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAAGCCTA
CTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTT
ACCTACTGTTGTTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCT
CTGGCAAAAAAAAAAAAAA

FIGURE 156

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885
```

```
><subunit 1 of 1, 432 aa, 1 stop
```

><MW: 47644, pI: 5.18, NX(S/T): 2

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYF
LCGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWF
SACFDNFTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSL
HCLACGKSLKTPRVVGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDV
FNWKVRAGSDKLGSPSLAVAKIIIIIEFNPMYPKDN DIALMKLQFPLTFSGTVRPICLPFFD
EELTPATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMACAGIPE
GGVDTQCQGDSSGGLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 157

GGGTGAGGACACTGAGACACCGGAAAGCCTGGCATTCCAGAGGGAGGGAAACGCAGCGGCATCCCCAGGCTCCAG
AGCTCCCTGGTGACAGTCTGTGGCTGAGCATGGCCCTCCCGACCCCTGGGCCTGGACCCCTGGAGCCTCCTGGGCC
TTTTCTCTTCCAAGTCTTCTGCTGCTGCTGCCGACGACGACCGCGGGGGAGGCGGGCAGGGGGCCCATGCCCA
GGGTGAGATACTATGCAGGGGATGAACGTAGGGCACTTAGCTTCTTCCACCAGAAGGGCCTCCAGGATTTTGACA
CTCTGCTCCTGAGTGGTGATGGAATACTCTCTACGTGGGGGCTCGAGAAGCCATTCTGGCCTTGGATATCCAGG
ATCCAGGGGTCCCAGGCTAAAGAACATGATACCGTGGCCAGCCAGTGACAGAAAAAGAGTGAATGTGCCTTTA
AGAAGAAGAGCAATGAGACACAGTGTTCAACTTCATCCGTGTCTGGTTTTCTTACAATGTACCCATCTCTACA
CCTGCGGCACCTTCGCCTTCAGCCCTGCTTGTACCTTCATTGAAGTTCAAGATTCTACCTGTTGCCCATCTCGG
AGGACAAGGTCATGGAGGGAAAAGGCCAAAGCCCCCTTTGACCCCGCTCACAAGCATACGGCTGTCTTGGTGGATG
GGATGCTCTATTCTGGTACTATGAACAACCTTCTGGGCAGTGAGCCCATCCTGATGCGCACACTGGGATCCCAGC
CTGTCTCAAGACCGACAACCTTCTCCGCTGGCTGCATCATGACGCCTCCTTTGTGGCAGCCATCCCTTCGACCC
AGGTCGTCTACTTCTTCTTCGAGGAGACAGCCAGCGAGTTTGACTTCTTTGAGAGGCTCCACACATCGCGGGTGG
CTAGAGTCTGCAAGAATGACGTGGGCGGCGAAAAGCTGCTGCAGAAGAAGTGGACCACCTTCTGAAGGCCCAGC
TGCTCTGCACCCAGCCGGGGCAGCTGCCCTTCAACGTATCCGCCACGCGGTCTGCTCCCCGCCGATTCTCCCA
CAGCTCCCCACATCTACGCAGTCTTCACCTCCAGTGGCAGGTTGGCGGGACCAGGAGCTCTGCGGTTTGTGCCT
TCTCTCTCTTGGACATTGAACGTGTCTTTAAGGGGAAATACAAAGAGTTGAACAAAGAACTTCACGCTGGACTA
CTTATAGGGGGCCCTGAGACCAACCCCCGGCCAGGCAGTTGCTCAGTGGGCCCTCCTCTGATAAGGCCCTGACCT
TCATGAAGGACCATTTCTGATGGATGAGCAAGTGGTGGGGACGCCCTGCTGGTGAAATCTGGCGTGGAGTATA
CACGGCTTGCAGTGGAGACAGCCAGGGCCTTGATGGGCACAGCCATCTTGTATGTACCTGGGAACCACCACAG
GGTCGCTCCACAAGGCTGTGGTAAGTGGGGACAGCAGTGTCTCATCTGGTGGAAGAGATTAGCTGTTCCCTGACC
CTGAACCTGTTGCAACCTGCAGCTGGCCCCCACCAGGGTGCAGTGTGTTGTAGGCTTCTCAGGAGGTGTCTGGA
GGGTGCCCCGAGCCAACTGTAGTGTCTATGAGAGCTGTGTGGACTGTGTCCTTGCCCGGGACCCCCACTGTGCCT
GGGACCTGAGTCCCGAACCTGTTGCCTCCTGTCTGCCCCAACCTGAACTCCTGGAAGCAGGACATGGAGCGGG
GGAACCCAGAGTGGGCATGTGCCAGTGGCCCCATGAGCAGGAGCCTTCGGCCTCAGAGCCGCCCGCAAATCATTA
AAGAAGTCTGGCTGTCCCAACTCCATCCTGGAGTCCCCGCCCCACCTGTGAGCCTTGGCCTCTTATTATT
GGAGTCATGGCCCAGCAGCAGTCCCAGAAGCCTCTTCCACTGTCTACAATGGCTCCCTCTTGCTGATAGTGCAGG
ATGGAGTTGGGGGTCTCTACCAGTGTGGGCAACTGAGAATGGCTTTTCATACCCTGTGATCTCCTACTGGGTGG
ACAGCCAGGACCAGACCCTGGCCCTGGATCCTGAACTGGCAGGCATCCCCGGGAGCATGTGAAGGTCCCCTTGA
CCAGGGTCAGTGGTGGGGCCGCCCTGGCTGCCCAGCAGTCTACTGGCCCCACTTTGTCACTGTCACTGTCTCT
TTGCCTTAGTGCTTTTCAGGAGCCCTCATCATCCTCGTGGCCTCCCCATTGAGAGCACTCCGGGCTCGGGGCAAGG
TTCAGGGCTGTGAGACCCTGCGCCCTGGGGAGAAGGCCCGTTAAGCAGAGAGCAACACCTCCAGTCTCCCAAGG
AATGCAGGACCTCTGCCAGTGTGTGGACGCTGACAACAACCTGCCTAGGCACTGAGGTAGCTTAAACTCTAGGCA
CAGGCCGGGGCTGCGGTGCAGGCACCTGGCCATGCTGGCTGGGCGGCCCAAGCACAGCCCTGACTAGGATGACAG
CAGCACAAAAGACCACCTTTCTCCCTGAGAGGAGCTTCTGCTACTCTGCATCACTGATGACACTCAGCAGGGTG
ATGCACAGCAGTCTGCCTCCCTATGGGACTCCCTTCTACCAAGCACATGAGCTCTCTAACAGGGTGGGGGCTAC
CCCCAGACCTGCTCCTACACTGATATTGAAGAACCTGGAGAGGATCCTTCAGTTCTGGCCATTCCAGGGACCCTC
CAGAAACACAGTGTTCAGAGACCCTAAAAAACCTGCCTGTCCAGGACCCTATGGTAATGAACACCAAACATC
TAAACAATCATATGCTAACATGCCACTCCTGGAACTCCACTCTGAAGCTGCCGCTTTGGACACCAACACTCCCT
TCTCCCAGGGTCATGCAGGGATCTGCTCCCTCCTGCTTCCCTTACCAGTCTGTCACCGCTGACTCCCAGGAAGTC
TTTCCTGAAGTCTGACCACCTTTCTTCTTGGCTTCAGTTGGGGCAGACTCTGATCCCTTCTGCCCTGGCAGAATGG
CAGGGGTAATCTGAGCCTTCTTCACTCCTTTACCCTAGCTGACCCCTTCACCTCTCCCCCTCCCTTTTCCTTTGT
TTTGGGATTTCAGAAAACCTGCTTGTGAGAGACTGTTTATTTTTTATTAAAAATATAAGGCTTAAAAAAA

FIGURE 158

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71166

><subunit 1 of 1, 761 aa, 1 stop

><MW: 83574, pI: 6.78, NX(S/T): 4

MALPALGLDPWSLLGLFLFQLLQLLLPPTTTAGGGGQGPMPRVRYAGDERRALSFFHQKGLQ
DFDTLLLSGDGNTLYVGAREAILALDIQDPGVPRCLKNMIPWPASDRKKSECAFKKKSNETQC
FNFIRVLVSYNVTHLYTCGTFAFSPACTFIELQDSYLLPISEDKVMEGKGQSPFDPAHKHTA
VLVDGMLYSGTMNNFLGSEPILMRTLGSQPVLKTDNFLRWLHHDASFVAAIPSTQVVYFFFE
ETASEFDDFERLHTSRVARVCKNDVGGEKLLQKKWTTFLKAQLLCTQPGQLPFNVIRHAVLL
PADSPTAPHIYAVFTSQWQVGGTRSSAVCAFSLLDIERVFKGKYKELNKETSRWTTYRGPET
NPRPGSCSVGPSSDKALTFMKDHFLMDEQVVGTPLLVKSGVEYTRLAVETAQGLDGHSHLVM
YLGTTTGSLLHKAVVSGDSSAHLVEEIQLFDPPEPVRNLQLAPTQGAVFVGFSGGVWRVPRAN
CSVYESCVCVLARDPHCAWDPESTRCCLLSAPNLNSWKQDMERGNPEWACASGPMSRSLRP
QSRPQIIKEVLAVPNSILELPCPHLSALASYWSHGPAAVPEASSTVYNGSLLLIVQDGVGG
LYQCWATENGFSYPVISYWVDSQDQTLALDPELAGIPREHVKVPLTRVSGGAALAAQQSYWP
HFVTVTVLFLVLSGALIILVASPLRALRARGKVQGCETLRPGEKAPLSREQHLQSPKECRT
SASDVDADNNCLGTEVA

Signal peptide:

amino acids 1-30

Transmembrane domains:

amino acids 136-156, 222-247, 474-490, 685-704

FIGURE 159

AGGGTCCCTTAGCCGGGCGCAGGGGCGCGCAGCCAGGCTGAGATCCGCGGGCTTCGGTAGAAG
TGAGC**ATGG**GCTGGGCAGCGAGTGCTTCTTCTAGTGGGCTTCCTTCTCCCTGGGGTCCCTGCTC
TCAGAGGCTGCCAAAATCCTGACAATATCTACAGTAGGTGGAAGCCATTATCTACTGATGGA
CCGGGTTTCTCAGATTCTTCAAGATCACGGTCATAATGTCACCATGCTTAACCACAAAAGAG
GTCCTTTTATGCCAGATTTTAAAAAGGAAGAAAAATCATATCAAGTTATCAGTTGGCTTGCA
CCTGAAGATCATCAAAGAGAATTTAAAAAGAGTTTTGATTTCTTTCTGGAAGAACTTTAGG
TGGCAGAGGAAAATTTGAAAACCTATTAAATGTTCTAGAATACTTGGCGTTGCAGTGCAGT
ATTTTTTAAATAGAAAGGATATCATGGATTCCTTAAAGAATGAGAACTTCGACATGGTGATA
GTTGAAACTTTTGACTACTGTCCTTTCCCTGATTGCTGAGAAGCTTGGAAGCCATTTGTGGC
CATTCTTTCCACTTCATTTCGGCTCTTTGGAATTTGGGCTACCAATCCCCTTGTCTTATGTT
CAGTATTCGGTTCCTTGCTGACTGATCACATGGACTTCTGGGGCCGAGTGAAGAATTTTCTG
ATGTTCTTTAGTTTCTGCAGGAGGCAACAGCACATGCAGTCTACATTTGACAACACCATCAA
GGAACATTTACAGAAGGCTCTAGGCCAGTTTTTGTCTCATCTTCTACTGAAAGCAGAGTTGT
GGTTCATTAACTCTGACTTTGCCTTTGATTTTGCTCGACCTCTGCTTCCCAACACTGTTTAT
GTTGGAGGCTTGATGGAAAACCTATTAAACAGTACCACAAGACTTGGAAGAACTTCATTGC
CAAGTTTGGGGACTCTGGTTTTGTCTTGTGACCTTGGGCTCCATGGTGAACACCTGTCAGA
ATCCGGAAATCTTCAAGGAGATGAACAATGCCTTTGCTCACCTACCCCAAGGGGTGATATGG
AAGTGTCAGTGTTCTCATTGGCCCAAAGATGTCCACCTGGCTGCAAATGTGAAAATTGTGGA
CTGGCTTCCTCAGAGTGACCTCCTGGCTCACCCAAGCATCCGTCTGTTTGTCACCCACGGCG
GGCAGAATAGCATAATGGAGGCCATCCAGCATGGTGTGCCCATGGTGGGGATCCCTCTCTTT
GGAGACCAGCCTGAAAACATGGTCCGAGTAGAAGCCAAAAGTTTGGTGTTTCTATTAGTT
AAAGAAGCTCAAGGCAGAGACATTGGCTCTTAAGATGAAACAAATCATGGAAGACAAGAGAT
ACAAGTCCGCGGCAGTGGCTGCCAGTGTATCCTCGCTCCACCCGCTCAGCCCCACACAG
CGGCTGGTGGGCTGGATTGACCACGTCTCCAGACAGGGGGCGCAGCACCTCAAGCCCTA
TGTCTTTCAGCAGCCCTGGCATGAGCAGTACCTGTTTCGACGTTTTTGTGTTTCTGCTGGGGC
TCACTCTGGGGACTCTATGGCTTTGTGGGAAGCTGCTGGGCATGGCTGTCTGGTGGCTGCGT
GGGGCCAGAAAGGTGAAGGAGACA**TAA**AGGCCAGGTGCAGCCTTGGCGGGGTCTGTTTGGTGG
GCGATGTCACCATTTCTAGGGAGCTTCCCACTAGTTCTGGCAGCCCCATTCTCTAGTCCTTC
TAGTTATCTCCTGTTTTCTTGAAGAACAGGAAAAATGGCCAAAATCATCCTTTCCACTTGC
TAATTTTGGCTACAAATTCATCCTTACTAGCTCCTGCCTGCTAGCAGAAATCTTCCAGTCCT
CTTGTCTCTCTTTGTTTGCCATCAGCAAGGGCTATGCTGTGATTCTGTCTCTGAGTGACTTG
GACCACTGACCCTCAGATTTCCAGCCTTAAATCCACCTTCTCTCATGCGCCTCTCCGAA
TCACACCCTGACTCTTCCAGCCTCCATGTCCAGACCTAGTCAGCCTCTCTCACTCCTGCCCC
TACTATCTATCATGGAATAACATCCAAGAAAGACACCTTGCATATTCTTTAGTTTCTGTTT
TGTTCTCCACATATTCTCTTCAATGCTCAGGAAGCCTGCCCTGTGCTTGAGAGTTTCAGGGC
CGGACACAGGCTCACAGGTCTCCACATTGGGTCCCTGTCTCTGGTGGCCACAGTGAGCTCCT
TCTTGGCTGAGCAGGCATGGAGACTGTAGTTTCCAGATTTCTTGAAAAATAAAAGTTTACA
GCGTTATCTCTCCCCAACCTCACTAA

FIGURE 160

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71169

><subunit 1 of 1, 523 aa, 1 stop

><MW: 59581, pI: 8.68, NX(S/T): 1

MAGQRVLLLVGFLLPGVLLSEAAKILTISTVGGSHYLLMDRVSQILQDHGHNVTMLNHKRGPFMPDFKKEEKSYQVISWLAPEDHQREFKKSFDFFLEETLGGRGKFENLLNVLEYLALQCSHF
LNRKDIMDSLKNENFDMVIVETFDYCPFLIAEKLGKPFVAILSTSFGSLEFGLPIPLSYVPVFRSLLTDHMDFWGRVKNFLMFFSFCRRQQHMQSTFDNTIKEHFTEGSRPVLSHLLKAEWLF
INSDFAFDFARPLLNTVYVGGLEMKPIKPVPQDLENFIAKFGDSGFVLVTLGSMVNTCQNP
EIFKEMNNAFAHLPQGVWKCQCShWPKDVHLAANKIVDWLPQSDLLAHPSIRLFVTHGGQ
NSIMEAIQHGVPMVGIPFGDQPENMVRVEAKKFGVSIQLKKLKAETLALKMKQIMEDKRYK
SAAVAASVILRSHPLSPTQRLVGWIDHVLQTGGATHLKPYPVQQPWHEQYLFDFVFLGLT
LGTWLWLCGKLLGMAVWWLRGARKVKET

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 483-504

GGGCTGTTGATTGTGTTGGGGGATTTTGAAGAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCT
GCCTCTGGCATATGCACACACTCACACATTCTGTACACCCGTCACACACACATAACCATGTT
CTCCATCCCCCAGGTCCAGCCCTCAGTGCTGTCCCATCCAGCAGGGGCTACCCTGAAGCTCT
GGCTGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCTTTCTCTCCCAAAGCCCCA
ACTGCTGTCACTGCATGCTCTGCCAAGGAGGAGGGAACTGCAGTGACAGCAGGAGTAAGAGT
GGGAGGCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTTCAGCGAGCCTAGAGAGGGC
AGACTATCAGGGGTGCCGGCGGTGAGAATCCAGGGAGAGGAGCGGAAACAGAAAGAGGGGGCAG
AGACCGGGGCACCTTGTGGGTTGCAGAGCCCCCTCAGCCATGTTGGGAGCCAAGCCACACTGGC
TACCAGGTCCCCTACACAGTCCCGGGCTGCCCTTGGTTCTGGTGCTTCTGGCCCTGGGGGCC
GGGTGGGCCCAGGAGGGGTTCAGAGCCCGTCCCTGCTGGAGGGGGAGTGCCTGGTGGTCTGTGA
GCCTGGCCGAGCTGCTGCAGGGGGGGCCCGGGGAGCAGCCCTGGGAGAGGCACCCCTGGGC
GAGTGGCATTTGCTGCGGTCCGAAGCCACCACCATGAGCCAGCAGGGGAAACCGGCAATGGC
ACCAGTGGGGCCATCTACTTCGACCAGGTCCCTGGTGAACGAGGGCGGTGGCTTTGACCGGGC
CTCTGGCTCCTTCGTAGCCCCTGTCCGGGGTGCTCTACAGCTTCGGGTCCATGTGGTGAAGG
TGTACAACCGCCAAACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCTCTCAGCC
TTTGCCAATGATCCTGACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGGA
CCCTGGGGACCGAGTGTCTCTGCGCCTGCGTCGGGGGAATCTACTGGGTGGTTGGAAATACT
CAAGTTTCTCTGGCTTCCTCATCTTCCCTCTCTCGAGGACCCAAGTCTTTC AAGCACAGAAT
CCAGCCCCTGACAACCTTTCTTCTGCCCTCTCTTGCCCCAGAAACAGCAGAGGCAGGAGAGAG
ACTCCCTCTGGCTCCTATCCACCTCTTTGCATGGGACCTGTGCCAAACACCCAAGTTTAA
GAGAAGAGTAGAGCTGTGGCATCTCCAGACCAGGCCTTTCCACCCACCCACCCCAAGTTACC
CTCCCAGCCACCTGCTGCATCTGTTCCCTGCCTGCAGCCCTAGGATCAGGGCAAGGTTTGGCA
AGAAGGAAGATCTGCACTACTTTGCGGCCTCTGCTCCTCCGTTCCCCCACCCAGCTTCCT
GCTCAATGCTGATCAGGGACAGGTGGCGCAGGTGAGCCTGACAGGCCCCCCACAGGAGCCAG
ATGGACAAGCCTCAGCGTACCCTGCAGGCTTCTTCCTGTGAGGAAAGCCAGCATCACGGATC
TCAGCCAGCACCGTCAGAAAGCTGAGCCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCAC
AGGCAGAAGGGTGGGAAGGGCCTGGAGTCTGTGGCTGGTGAGGAAGGAAGGAGGGTGTATTG
TCTAGACTGAACATGGTACACATTCTGCATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCT
GGCTGTCCTTCTATGCTGGATCCCAGATGGACTCTGGCCCTTACCTCCCCACCTGAGATTAG
GGTGAGTGTGTTTGTCTCTGGCTGAGAGCAGAGCTGAGAGCAGGTATACAGAGCTGGAAGTGG
ACCATGGAAAACATCGATAACCATGCATCCTCTTGCTTGGCCACCTCCTGAAACTGCTCCAC
CTTTGAAGTTTGAACCTTAGTCCCTCCACACTCTGACTGCTGCCTCCTTCCTCCCAGCTCTC
TCACTGAGTTATCTTCACTGTACCTGTTCCAGCATATCCCCACTATCTCTCTTTCTCCTGAT
CTGTGCTGTCTTATTCTCCTCCTTAGGCTTCCTATTACCTGGGATTCCATGATTCATTCTT
CAGACCTCTCCTGCCAGTATGCTAAACCTCCCTCTCTCTTTCTTATCCCGCTGTCCCATT
GGCCCAGCCTGGATGAATCTATCAATAAAACAAGTAGAGAATGGTGGTCAGTGAGACACTAT
AGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGGATCGGGTGTTACAGGTACAAGTAGGTA
TGTTGCAGAGGAAAATAAATATCAAAGTGTATACTAAATTAATAA

FIGURE 162

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71180
```

```
><subunit 1 of 1, 205 aa, 1 stop
```

><MW: 21521, pI: 7.07, NX(S/T): 1

MLGAKPHWLPGLHSPGLPLVLVLLALGAGWAQEGSEPVLLERGECLVVCPEGRAAGGPGGA
ALGEAPPGRVAFAAVRSHHHEPAGETGNGTSGAIYFDQVLVNEGGGFDRASGSFVAPVRGVY
SFRFHVVVKVYNRQTVQVSLMLNTWPVISAFAANDPDVTTREAAATSSVLLPLDPGDRVSLRLRG
NLLGGWKYSSFSGFLIFPL

Signal peptide:

amino acids 1-32

FIGURE 163

[illegible]

FIGURE 164

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184
><subunit 1 of 1, 388 aa, 1 stop
><MW: 43831, pI: 9.64, NX(S/T): 3
MKTLLAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRS
KVEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLI AVL YFTWL VFDWNTPKKGGRS
QWVRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAF CNFSTEATEVSKKFP
GIRPYLATLAGNFRMPVLREYLMSGGICPVS RDTIDYLLSKNGSGNAIIIVVGGAAESLSSM
PGKNAVTLRNRKGFVKLALRHGADLVPIYSEFGENEVYQVIFEEGSWGRWVQKKFQKYIGFA
PCIFHGRGLFSSDTWGLVPYSKPITTVVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDK
HKTKFGLPETEVLEVN

Important features of the protein:**Transmembrane domain:**

amino acids 76-97

N-glycosylation sites.

amino acids 60-63, 173-176, 228-231

N-myristoylation sites.

amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-245, 318-323, 378-383

FIGURE 165

GGGCGGCGGGATGGGGGCCGGGGCGGGCGGGCGCCGCACTCGCTGAGGCCCCGACGCAGGGCCGGGCGGGCCCA
 GGGCCGAGGAGCGCGGCGGCCAGAGCGGGGCGCGGAGGCGACGCCGGGGACGCCCGCGCGACGAGCAGGTGGCG
 GCGGCTGCAGGCTTGTCCAGCCGGAAGCCCTGAGGGCAGCTGTTCCCACTGGCTCTGCTGACCTTGTGCCTTGA
 CGGCTGTCCTCAGCGAGGGGCGCTGCACCCGCTCCTGAGCAGCGCCATGGGCCTGCTGGCCTTCTGAAGACCCA
 GTTCGTGCTGCACCTGCTGGTCGGCTTTGTCTTCGTGGTGAGTGGTCTGGTCATCAACTTCGTCCAGCTGTGCAC
 GCTGGCGCTCTGGCCGGTCAGCAAGCAGCTCTACCGCCGCTCAACTGCCGCTCGCCTACTCACTCTGGAGCCA
 ACTGGTCATGCTGCTGGAGTGGTGGTCTGCACGGAGTGACACTGTTACGGACCAGGCCACGGTAGAGCGCTT
 TGGGAAGGAGCACGCAGTCATCATCTCAACCACAACCTCGAGATCGACTTCTCTGTGGGTGGACCATGTGTGA
 GCGCTTCGGAGTGTGGGGAGCTCCAAGTCTCGCTAAGAAGGAGCTGCTCTACGTGCCCCCTCATCGGCTGGAC
 GTGGTACTTTCTGGAGATTGTGTTCTGCAAGCGGAAGTGGGAGGAGGACCGGGACACCGTGGTGAAGGGCTGAG
 GCGCCTGTGCGACTACCCCGAGTACATGTGGTTTCTCCTGTACTGCGAGGGGACGCGCTTCACGGAGACCAAGCA
 CCGCGTTAGCATGGAGGTGGCGGCTGCTAAGGGGCTTCTGTCTCAAGTACCACCTGCTGCCGCGGACCAAGGG
 CTTCAACCACCGCAGTCAAGTGCCTCCGGGGGACAGTCGCAGCTGTCTATGATGTAACCCTGAACTTCAGAGGAAA
 CAAGAACCCGTCCCTGCTGGGGATCCTCTACGGGAAGAAGTACGAGGCGGACATGTGCGTGAGGAGATTTCTCT
 GGAAGACATCCCGCTGGATGAAAAGGAAGCAGCTCAGTGGCTTCATAAACTGTACCAGGAGAAGGACGCGCTCCA
 GGAGATATATAATCAGAAGGGCATGTTTCCAGGGGAGCAGTTTAAGCCTGCCCGGAGGCCGTGGACCCTCTGAA
 CTTCTGTCTGGGCCACCATTCTCCTGTCTCCCTCTTCAGTTTTGTCTTGGGCGTCTTTGCCAGCGGATCACC
 TCTCCTGATCCTGACTTTCCTTGGGGTTTGTGGGAGCAGCTTCCTTTGGAGTTTCGCAGACTGATAGGAGAATCGCT
 TGAACCTGGGAGGTGGAGATTGCAGTGAAGCTGAGATGGCATCACTGTACTCCAGCCTAGGCAACAGAGCAAGACT
 CAGTCTCAAAAAAAAAAAAAAAAAACAAAAACCCAGAAATTCTGGAGTTGAACTGTGTAGTTACTGACATGAAAA
 ATTCACTAGAGGCTGAACAGCAGATTTGAGCAGGCAGAAAAAATCAGCAAGCTTGAAGATGGTACCTTGAGATT
 TTTCAAGGCTAATGAAAAAGAATGAAGGAAAATTAACAGCCTCAGAGACCCATGGTGCACCGTCACACAAATCAA
 CATATGCATGATGAGAGTCCCAGAAGGAGAGGAGAGAAAGGGTCAGAAAGAATGGCCACAAGCTGATGAAAAACA
 GTAACCTACCCACTCAGGAAGCTCAGTGAACCTCAATGAGGATGAATATCAGAGATCCACACCTAGATATTTAT
 AATCAAAGTGTCAAATGACAAAGAATCTTGAAAGCAGCAAGAGATGAGCAACTTATCTTGTTCAAAGGATCTTTG
 ATCAGATTAACAGCTCATTTCTCCTCAGAAATCATGGGAGCCAGGAGATAGTGGGATGAACACTGTTGAAGGCAA
 AACCTTCAACTGTAATTATTGGACTTTTGTAGTCTTAGATGGTCTGACCTCTTTGTCTTCAGGGACAGTTTTTCA
 ATTTAATCCCTAATAACAATTAGTCAAGCTTCCTTGACCTGTAGGAAGGCCTGTCTTTAGGCCGGGCACAGTGGC
 TTACACCTGTAATCCCAGCACTTTGGGAGGCCAGACGGGTGGATCATTTGGGGTCAGGCTGATCTCAAACCTCT
 GAGTTCAGGTGATCTGCCCCCTCAGCCTCCCAAAGTGTTGTGATTGCAGGCGTGAGCCACTGCGCCTGGCCGGA
 ATTTCTTTTAAAGGCTGAATGATGGGGGCCAGGCACGATGGCTCACGCCTGTGATCCCAAGTAGCTTGGATTGTA
 AACATGCACCACCATGCCTGGCTAATTTTTGTATTTTTTAGTAGAGACGTGTTAGCCAGGCTGGTCTCGATCTCCT
 GACCTCAAGTGACCACCTGCCTCAGCCTCCCAAAGTACTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTTGA
 GCATCTTGTGATGTGCTTATTGGCCATTTGTATATCTTCTATCTTCTTTGGGGAAATGTCTGTTCAAGTCCTTTG
 CCTTTTTAAATTTTTATTATTTATTTATTTATTTATTTTGTAGACAGGGTCTTGTCTGTTGCCAGGCTGGAGTA
 CAGTGGCACAGTCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTGCAGTGATCCTCCACCTCAGCCTCCCTTGT
 AGCTGTATTTTTTTGTATTTTGTATTTTGTAGCTGTAGTTTTTGTATTTTTTGTGGAGACAGCATTTACCATGA
 TGCCCAGGCTGGTCTTGAACCTCTGAGCTCAAGTGATCGCTGCTTCAGCCTCCCAAAGTGCTGGGATTACAGA
 CATGAGCCACTGCACCTGGCAAACCTCCCAAATTAACACACACACACAAAAAACACCTGATTCAAAATGGGCA
 GAGGGGCCGGGTGTGGCCCCAACTACCAGGGAGACTGAAGTGGGAGGATCGCTTGGGCATGAGAAGTCGAGGCTG
 CAGTGAGTCGAGGTTGTGCGACTGCATTCCAGCCTGGACAACAGAGTGAGACCCTGTCTC

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 166

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213
><subunit 1 of 1, 368 aa, 1 stop
><MW: 42550, pI: 9.11, NX(S/T): 1
MGLLAFLKTQFVLHLLVGFVFVVSGLVINQVQLCTLALWPVSKQLYRRLNCRSLAYSLWSQLV
MLEWWSCTECTLFTDQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGVLGSSKVLAKK
ELLYVPLIGWTWYFLEIVFCKRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHR
VSMEVAAAKGLPVLKYHLLPRTKGFTTAVKCLRGTVAAYVDVTLNFRGNKNPSLLGILYGKK
YEADMCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLN
FLSWATILLSPLFSFVLGVFASGSPLLILTFLGFVGAASFGVRRLIGESLEPGRWRLQ
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

Transmembrane domains:

amino acids 307-323, 335-352

Tyrosine kinase phosphorylation sites.

amino acids 160-168, 161-169

FIGURE 167

GATATTCTTTATTTTAAAGAATCTGAAGTACTATGCATCACTCCCTCCAATGTCCTGGGGCA
 GCCACCAGGCATATTCATCTTTGTGTGTGTTTTCTTTGCTTTAGCACTGGGGCACTTCTT
 GCTTATTTCTTTGGTAGGAAAGGGGCTCAGTTTGTCTTGTGGGGTTGGTGGCAGGCAGGCCG
 GCTTACGCCTGATACGGCCCTGGGTTAGAAGGGAAGGGAAGATAAACTTTTATACAAATGGG
 GATAGCTGGGGTCTGAGACCTGCTTCCTCAGTAAAATTCCTGGGATCTGCCTATACCTTCTT
 TTCTCTAACCTGGCATAACCCTGCTTAAAGCCTCTCAGGGCTTCTCTCTGTTCTTAGGATCAA
 AGTATTTAGAGCTACAAGAGCCCTCATGGTCTGGCCCCCTGCCCCCTGGCCAGCTTCATTGT
 ACATGTGGTGTCTCTTGTCTGTTCCCTGTAATGTGGTATGCCATGGGGTCTTTGCACAAGCCT
 TTCCTCTTTGGCTGGACACTGTTCCCTGCCCCCCCCATACTCTTCCTACTTAATATGTAGTC
 ATCCTGCAGATTTCAATTCTAACATCATTTTCTCCAGGGATCCTGGCCTGACAGAATCTCAT
 CTTGTTTAATGCTCTCATAAGACCACTTGTTTTCCCTTTTGCAGCACTTGCCACTCAGTTGTA
 TCTTTATGTGCGTTTGTGGTTGTATGGGTGTGTCTGTTCCCCAGAATGCCCAGCTCTGAGC
 TCGGTGAGGGTCAAGGGCATTGCTGTGCCTGCCAGGTATAGTGCCTACATGTGGTGGGTGCT
 CATGTTTTAGAGACTAAATGGAGGAGGAGATGAGGAAAAGATTGAAATCTCTCAGTTCACCA
 GATGGTGTAGGGCCCAGCATTGTAAATTCACACGTTGACTGTGCTTGTGAATTATCTGGGGA
 TGCAGGTCCTGATTCAGTAGGCCCCAGGTTGGGCATCTCTAACAACTCCCACGTGATGCTGA
 TGCTGGTCCTATGAACTATACTAAATAGTAAGAATCTATGGAGCCAGGCTGGGCATGGTGGC
 TCACACCTATGATCCCAGCACTTTGGGAGGCTGAGGCAGGCTGATCACCTGGAGTCAGGATT
 TCAAGACTAGCCTGGCCAACATGGTGGAACCCCATCTGTACTAAAAATACACAAATTAGCTG
 GGCATGGTGGCACATGCCTGTAGTCCCAGCTACTTGGGAGGCTGAAGCAAGAGAATCGCTTG
 AACCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCAGGCCACTGTATTCCAACCAGGGTGAC
 AGAGTGAGACTCTATGTCCAAAAAAAAAAAAA

FIGURE 168

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234
><subunit 1 of 1, 143 aa, 1 stop
><MW: 15624, pI: 9.58, NX(S/T): 0
MHHS LQCPGAATRHIHLCVCF SFALALGHFLLISLVGKGLSLSCGVGGRQAGLRLIRPWVRR
EGKINFYTNGDSWGLRPASSVKFLGSAYTFFSLTWHTLLKASQGFS LFLGSKYLELQEPSWS
GPCPPGQLHCTCGVLLSFL
```

Important features of the protein:**Signal peptide:**

amino acids 1-28

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 169

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGA
 TTATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCAC
 AGGAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTAC
 ATTAAAATCTGTTTTTTTGTCTCTTGTAAGTACCTTTACCTTCCTAACACAGAGGATCTGT
 CACTGTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCAC
 ACCGTCCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTC
 ACCAACTGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGACAGGTAGCTGAGCCTCTTGGA
 GCTGCGGCTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTT
 CATAGGCGATGGCTCCCACTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGG
 CCAGGACGGGCGGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCCGCCAT
 CTAACCTTTTCATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGG
 AACCCAGAGCTTGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTT
 TATGTGACAGGACTTGCAATCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCA
 GGGAAGGAACTTGTGCCAAATTATGGGTGAGAAAAGATGGAGGTGTTGGGTATCACAAGGC
 ATCGAGTCTCCTGCATTCAGTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACT
 CGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCGATCCACGTACCAGCTGCTG
 AAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAG
 CCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCCTCCTTCCCTC
 TGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATG
 GCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCG
 AACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCT
 CAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGT
 TCAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAA
 TCCAAACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATG
 CCAACATTTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTG
 GCCAACATGGTGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGT
 GTGTGCCTGTATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGT
 GAAGGAGGCTGAGACAGGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGA
 AAAAATAAAAAAAGAATTATGGTTATTTGTAA

FIGURE 170

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71277
```

```
><subunit 1 of 1, 109 aa, 1 stop
```

><MW: 11822, pI: 8.63, NX(S/T): 0

MLWWLVLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTA

SPCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

CGGGGGCCGCGAGTCCGAGACCTGTCCCAGGAGCTCCAGCTCACGTGACCTGTCACTGCCTC
CCGCCGCCTCCTGCCCGCGCCATGACCCAGCCGGTGCCCCGGCTCTCCGTGCCCGCCGCGCT
GGCCCTGGGCTCAGCCGCACTGGGCGCCGCCTTCGCCACTGGCCTCTTCCTGGGGAGGCGGT
GCCCCCATGGCGAGGCCGCGAGAGCAGTGCCTGCTTCCCCCGAGGACAGCCGCCTGTGG
CAGTATCTTCTGAGCCGCTCCATGCGGGAGCACCCGGCGCTGCGAAGCCTGAGGCTGCTGAC
CCTGGAGCAGCCGCAGGGGGATTCTATGATGACCTGCGAGCAGGCCCAGCTCTTGGCCAACC
TGGCGCGGCTCATCCAGGCCAAGAAGGCGCTGGACCTGGGCACCTTCACGGGCTACTCCGCC
CTGGCCCTGGCCCTGGCGCTGCCCGCGGACGGGCGCGTGTTGACCTGCGAGGTGGACGCGCA
GCCCCCGGAGCTGGGACGGCCCCCTGTGGAGGCAGGCCGAGGCGGAGCACAAAGATCGACCTCC
GGCTGAAGCCCGCCTTGAGACCCCTGGACGAGCTGCTGGCGGCGGGCGAGGCCGGCACCTTC
GACGTGGCCGTGGTGGATGCGGACAAGGAGAACTGCTCCGCCTACTACGAGCGCTGCCTGCA
GCTGCTGCGACCCGGAGGCATCCTCGCCGTCTCAGAGTCTGTGGCGCGGGAAGGTGCTGC
AACCTCCGAAAGGGGACGTGGCGGCCGAGTGTGTGCGAAACCTAAACGAACGCATCCGGCGG
GACGTCAGGGTCTACATCAGCCTCCTGCCCCTGGGCGATGGACTCACCTTGGCCTTCAAGAT
CTAGGGCTGGCCCCCTAGTGAGTGGGCTCGAGGGAGGGTTGCCTGGGAACCCAGGAATTGAC
CCTGAGTTTTTAAATTGAAAATAAAGTGGGGCTGGGACACAAAAAAAAAAAAAAAAAAAA

FIGURE 172

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71282
><subunit 1 of 1, 262 aa, 1 stop
><MW: 28809, pI: 8.80, NX(S/T): 1
MTQPVPRLSVPAALALGSAALGAAFATGLFLGRRCPPWRGRREQCLLPEDSRLWQYLLSRS
MREHPALRSLRLLTLEQPQGDSMMTCEQAQLLANLARLIQAKKALDLGTFTGYSALALALAL
PADGRVVTCEVDAQPPELGRPLWRQAEAEHKIDLRLKPALETLDELLAAGEAGTFDVAVVDA
DKENCSAYYERCLQLLRPGGILAVLRVLWRGKVLQPPKGDVAAECVRNLNERIRRDVRVYIS
LLPLGDGLTLAFKI
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

Transmembrane domains:

amino acids 8-30, 109-130

N-glycosylation site.

amino acids 190-193

Tyrosine kinase phosphorylation site.

amino acids 238-246

N-myristoylation sites.

amino acids 22-27, 28-33, 110-115, 205-210, 255-260

Amidation sites.

amino acids 31-34, 39-42

FIGURE 173

CCGCCGCCGACGCCGCTACCGCCGCTGCAGCCGCTTTCCGCGGCCCTGGGCTCTCGCCGTCA
GC**ATG**CCACACGCCTTCAAGCCCGGGGACTTGGTGTTCGCTAAGATGAAGGCTTACCCTCAC
TGGCCTGCCAGGATCGACGACATCGCGGATGGCGCCGTGAAGCCCCACCCAACAAGTACCC
CATCTTTTTTCTTTGGCACACACGAAACAGCCTTCCTGGGACCCAAGGACCTGTTCCCTACG
ACAAATGTAAAGACAAGTACGGGAAGCCCAACAAGAGGAAAGGCTTCAATGAAGGGCTGTGG
GAGATCCAGAACAACCCCCACGCCAGCTACAGCGCCCTCCGCCAGTGAGCTCCTCCGACAG
CGAGGCCCCCGAGGCCAACCCCGCCGACGGCAGTGACGCTGACGAGGACGATGAGGACCGGG
GGGTCAATGGCCGTACAGCGGTAACCGCCACAGCTGCCAGCGACAGGATGGAGAGCGACTCA
GACTCAGACAAGAGTAGCGACAACAGTGGCCTGAAGAGGAAGACGCCTGCGCTAAAGATGTC
GGTCTCGAAACGAGCCCGAAAGGCCTCCAGCGACCTGGATCAGGCCAGCGTGTCCCCATCCG
AAGAGGAGAACTCGGAAAGCTCATCTGAGTCGAGAAGACCAGCGACCAGGACTTCACACCT
GAGAAGAAAGCAGCGGTCCGGGCGCCACGGAGGGGGCCCTCTGGGGGGACGGAAAAAAGAA
GGCGCCGTACGCTCCGACTCCGACTCCAAGGCCGATTTCGGACGGGGCCAAGCCTGAGCCGG
TGGCCATGGCGCGGTTCGGCGTCTCCTCCTCCTCCTCCTCCTCCTCCTCCGACTCCGATGTG
TCTGTGAAGAAGCCTCCGAGGGGCGAGGAAGCCAGCGGAGAAGCCTCTCCCGAAGCCGCGAGG
GCGGAAACCGAAGCCTGAACGGCCTCCGTCCAGCTCCAGCAGTGACAGTGACAGCGACGAGG
TGGACCGCATCAGTGAGTGGAAGCGGCGGGACGAGGCGGGAGGCGCGAGCTGGAGGCCCGG
CGGCGGCGAGAGCAGGAGGAGGAGCTGCGGCGCTGCGGGAGCAGGAGAAGGAGGAGAAGGA
GCGGAGGCGCGAGCGGGCCGACCGCGGGGAGGCTGAGCGGGGCAGCGGCGGCAGCAGCGGGG
ACGAGCTCAGGGAGGACGATGAGCCCGTCAAGAAGCGGGGACGCAAGGGCCGGGGCCGGGGT
CCCCCGTCTCCTCTGACTCCGAGCCCGAGGCCGAGCTGGAGAGAGAGGCCAAGAAATCAGC
GAAGAAGCCGCAGTCTCAAGCACAGAGCCCGCCAGGAAACCTGGCCAGAAGGAGAAGAGAG
TGCGGCCCCGAGGAGAAGCAACAAGCCAAGCCCGTGAAGGTGGAGCGGACCCCGGAAGCGGTCC
GAGGGCTTCTCGATGGACAGGAAGGTAGAGAAGAAGAAAGACCCCTCCGTGGAGGAGAAGCT
GCAGAAGCTGCACAGTGAGATCAAGTTTGCCCTAAAGGTCGACAGCCCGGACGTGAAGAGT
GCCTGAATGCCCTAGAGGAGCTGGGAACCTGCAGGTGACCTCTCAGATCCTCCAGAAGAAC
ACAGACGTGGTGGCCACCTTGAAGAAGATTGCGCGTTACAAAGCGAACAAGGACGTAATGGA
GAAGGCAGCAGAAGTCTATACCCGGCTCAAGTTCGCGGGTCTCGGCCCAAAGATCGAGGCGG
TGCAGAAAGTGAACAAGGCTGGGATGGAGAAGGAGAAGGCCGAGGAGAAGCTGGCCGGGGAG
GAGCTGGCCGGGGAGGAGGCCCCCCCAGGAGAAGGCGGAGGACAAGCCCAGCACCGATCTCTC
AGCCCCAGTGAATGGCGAGGCCACATCACAGAAGGGGGAGAGCGCAGAGGACAAGGAGCACG
AGGAGGGTTCGGGACTCGGAGGAGGGGCCAAGGTGTGGCTCCTCTGAAGACCTGCACGACAGC
GTACGGGAGGGTCCCGACCTGGACAGGCCTGGGAGCGACCGGCAGGAGCGCGAGAGGGCAGC
GGGGGACTCGGAGGCCCTGGACGAGGAGAGCT**TGA**AGCCGCGGGCAGCCAGGCCCCAGCCCCCGC
CCGAGCTCAGGCTGCCCCCTCTCCTTCCCCGGCTCGCAGGAGAGCAGAGCAGAGAAGTGTGGG
GAACGCTGTGCTGTTTGTATTTGTTCCCTTGGGTTTTTTTTTCTGCCTAATTTCTGTGATT
TCCAACCAACATGAAATGACTATAAACGGTTTTTTTAATGA

FIGURE 174

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286
><subunit 1 of 1, 671 aa, 1 stop
><MW: 74317, pI: 7.61, NX(S/T): 0
MPHAFKPGDLVFAKMGYPHWPARIIDDIADGAVKPPPKNKYPIFFFGTHETAFLGPKDLFPYD
KCKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDDERD
VMAVTAVTATAASDRMESDSDSDKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSE
EENSESSSESEKTSQDFTPEKKA AVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPV
AMARSASSSSSSSSSSSDSDSVVKKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEV
DRISEWKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGD
ELREDDEPVKKRGRKGRGRGPPSSSDSEPEAELEAREAKKSAKKPQSSSTEPARKPGQKEKRV
RPEEKQQAQKPVKVERTRKRSEGFSMDRKVEKKKEPSVEEKLQKLHSEIKFALKVDS PDVKRC
LNALEELGTLQVTSQILQKNTDVVATLKKIRRYKANKDVMEKAAEVYTRLKSRVLGPKIEAV
QKVNKAGMEKEKAEEKLAGEELAGEEAPQEKAEDKPSTDLSAPVNGEATSQKGESAEDKEHE
EGRDSEEGPRCGSSEDLHDSVREGPDLDPRGSDRQERERARGDSEALDEES

Signal peptide:

amino acids 1-13

GTGTTGTTCTCCTGGATCTTCCACCTTACCAACTGCAGATCTTGGGACTCATCAGCCTCAATAATTATATTAATAATTA
ACACCATTTGAAAGAGAACATTGTTTTCATCATGAATGCTAATAAAGATGAAAGACTTAAAGCCAGAAGCCAAGA
TTTTACCTTTTTCTGCTTTGATGATGCTAAGCATGACCATGTTGTTTCTTCCAGTCACTGGCACTTTGAAGCA
AAATATTTCCAAGACTCAAGCTAACCTACAAAGACTTGCTGCTTTCAAATAGCTGTATTCCCTTTTTGGGTTCATC
AGAAGGACTGGATTTTCAAACCTCTTCTCTTATAGTGAGGAAAGAGGCAGGCTGCTCTTTGGGAGCCAAAGACCACAT
CTTTCTACTCAGTCTGGTTGACTTTAAACAAAATTTAAGAAGATTTATTGGCCTGCTGCAAAGGACCGGTGGA
ATTATGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAAATTTTCATCAGAGTACTTCAGCCCTATAACAA
AACTCACATATATGTGTGTGGAACCTGGAGCATTTTCATCCAATATGTGGGTATATTGATCTTGGAGTCTACAAGGA
GGATATTATATTCAAACCTAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCCTTTGCATCCTCAGCAGC
TTTTGTCTTCAGTAAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCCCTTGGCAAAGATACTGCATT
CACTCGATCCCTTTGGGCTACTCATGACCACCACACTACATCAGAAGTACATTTTCAGAGCATACTGGCTCAATGG
AGCAAAATTTATTGGAATTTTCTTATACCAGACACACATACATAACAGATGATGATAAAATATATTTCTTCTTCTG
TGAATCATCTCAAGAAGGCAGTACCTCCGATAAAACCATTCTTTCTCGAGTTGGAAGAGTTTGTAAAGATGATGT
AGGAGGACAACGCAGCCTGATAAAACAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTCCTGGAAG
TGATGGGGCAGATACTTACTTTTGATGAGCTTCAAGATATTTATTTACTCCCCACAAGAGATGAAAGAAATCCTGT
AGTATATGGAGTCTTTTACTACAACCAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTGTATAGCATGGCTGACAT
CAGAGCAGTTTTTAATGGTCCATATGCTCATAAGGAAAGTGCAGACCATCGTTGGGTGCAGTATGATGGGAGAA
TCCTTATCCACGGCCTGGTACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGA
TGATGTCATCAGTTTTCATAAAGCGGCACCTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGAGGCCAACGTT
CAAGAGAATCAATGTGGATTACAGACTGACACAGATAGTGGTGGATCATGTTCATTGCAGAAGATGGCCAGTACGA
TGTAATGTTTTCTTGGAACAGACATTGGAACCTGTCTCAAAGTTGTGTCAGCATTTCAAAGGAAAAGTGGAAATATGGA
AGAGGATAGTGTGGAGGAGTTGCAGATATCCAAGCACTCATCAATCATCTTGAACATGGAATTTGTCTCTGAAGCA
GCAACAATTTGATCATTGGTTCCGAGATGGATTAGTTGACGCTCTCCTTGCACAGATGCGACACTTATGGGAAAGC
TTGCGCAGACTGTTGTCTTGCCAGAGAGCCCTACTGTGCCTGGGATGGAAATGCATGCTCTCGATATGCTCCTAC
TTCTAAAAGGAGAGCTAGACGCCAAGATGTAAATATGGCGACCCAATCACCAGTGTCTGGGACATCGAAGACAG
CATTAGTCATGAAACTGCTGATGAAAAGGTGATTTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATACC
TAAATCCCAACAAGCAACTATTAAATGGTATATCCAGAGGTGAGGGGATGAGCATCGAGAGGAGTTGAAGCCCGA
TGAAGAATCATCAAACGGAATATGGGCTACTGATTTCGAAGTTTGCAGAAGAAGGATTCTGGGATGTATTACTG
CAAAGCCAGGAGCACACTTCTATCCACACCATGATGAAGTGCATCTTTGAATGTCAATTGAGAATGAACAGATGGA
AAATACCCAGAGGGCAGAGCATGAGGAGGGGCGAGGTCAAGGATCTATTGGCTGAGTCACGGTTGAGATCAAAAGA
CTACATCCAAATCCTTAGCAGCCCAAACCTTCAGCCTCGACCAGTACTGCGAACAGATGTGGCACAGGGAGAAAGCG
GAGACAGAGAAACAAGGGGGGGCCCAAAGTGGAAGCACATGCAGGAAATGAAGAAGAAACGAAATCGAAGACATCA
CAGAGAACCTGGATGAGCTCCCTAGAGCTGTAGCCACGTAGTTTTCTACTTAATTTAAAGAAAAGAATTCCTTACC
TATAAAACCATTGCCTTCTGTTTTGTATATCCCTTATAGATTAATTCATAAATGCTTTCCCATGGAGTTTTGCTAAGG
CACAAGACAATAATCTGAATAAGACAAATATGTGATGAATATAAGAAAGGCCAAAAAATTCATTGAAACGAGTTTT
CCAAGAACAATCTTGCACAAGCAAAGTATAAGAATTATCCTAAAAATAGGGGGTTTACAGTTGTAAATGTTTTTA
TGTTTTGAGTTTTTGAATTTTATTGTGATGATAAATAGTTGAGCTAAGCAAGCCCCGAATTTGATAGTGTATAAGGT
GCTTTATTTCCCTCGAATGTCCATTAAGCATGGAATTTACCATGCAGTTGTGCTATGTTCTTATGAACAGATATAT
CATTCCTATTGGAACACAGCTACCTTTGTGGTAGGGAATAAGAGGTGAGACACAAATTAAGACAACATCCCATTATC
AACAGGAACTTTCTCAGTGAGCCATTCACTCTGGAGAATGGTATAGGAATTTGGAGAGGTGCATTATTTCTTCT
TGGCCACTGGGGTTAAATTTAGTGTACTACAACATTTGATTTACTGAAGGGCACTAATGTTTTCCCCGAGATTTCT
ATTGACTAGTCAGGAGTAACAGGTTACAGAGAGAAGTTGGTGCTTAGTTATGTGTTTTTTAGAGTATATACTAA
GCTCTACAGGGACAGAAATGCTTAATAAATACTTTAATAAGATATGGGAAAATATTTAATAAAACAAGGAAAACA
TAATGATGTATAATGCATCCTGATGGGAAGGCATGCAGATGGGATTTGTTAGAAGACAGAAGGAAAGACAGCCAT
AAATCTGCTGTTTTGGGAAAACATCATATCCCCATGAAAGAAAGAACAATACAAAATAAGTGAGAGTAATGTAA
TGGAGCTCTTTTCACTAGGGTATAAGTAGCTGCCAATTTGTAATTCATCTGTGTAAGAAAATCTAGATTATAACA
AACTGCTAGCAAAATCTGAGGAAACATAAATTCCTTCTGAAGAATCATAGGAAGAGTAGACATTTTATTATAACC
AATGATATTTTCACTATATATTTTCTCTCTTTTAAAAAATATTTATCATACTCTGTATATTATTTCTTTTTACTGC
CTTTATTTCTCTCTGTATATTTGGATTTTGTGATTATATTTGAGTGAATAGGAGAAAACAATATATAACACACAGA
GAATTAAGAAAATGACATTTCTGGGGAGTGGGGATATATATTTGTTGAATAACAGAACGAGTGTAAATTTTAAAC
AACGGAAGAGGGTTAAATTTAACTTTTACATCTTCTGACATCTTCACTCAACCTTTTTCTCATTGCTGAGTTAATCTGTTGTAATT
GTAGTATTGTTTTGTAATTTAACAATAAATAAGCCTGCTACATGT

FIGURE 176

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883
 ><subunit 1 of 1, 777 aa, 1 stop
 ><MW: 89651, pI: 7.97, NX(S/T): 3
 MNANKDERLKARSQDFHLFPA LMMLSM TMLFLPVTGTLKQNI PRLKLT YKD LLLS NSCIPFL
 GSSEGLDFQTLLLD EERGRLLL GAKDHIFLLSLVDLNKNFKKIYWPA AKERVELCKLAGKDA
 NTECANFIRVLQPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFD
 PQQPFASVMTDEYLYSGTASDFLGKDTAFT RSLGPTH DHHYIRTDISEHYWLN GAKFIGTFF
 IPDTYNPDDDKIYFFFRESSQEGSTDKTILSRVGRVCKNDVGGQ RSLINKWTTFLKARLIC
 SIPGSDGADTYFDELQDIYLLPTRDERNPVVYGVFTTTSSIFKGS AVCVYSMADIRAVFNGP
 YAHKESADHRWVQYDGRIPYPRPGTCPSKTYDPLIKSTRDFPD DVISFIKRHSV MYKSVYPV
 AGGPTEFKRINVDYRLTQIVVDHVIAEDGQYDVMFLGTDIGTVLKVVSISKEKWNMEEVVLEE
 LQIFKHSSIIILNMELSLKQQQLYIGSRDGLVQLSLHRCDTY GKACADCC LARDPYCAWDGNA
 CSRYAPTSKRRARRQDVKYGDPITQCWDIEDSISHETADEKVI FGI EFNSTFLECIPKSQQA
 TIKWYIQRSGDEHREELKPDERIIKTEYGLLIRSLQKKDSGMY YCKAQEHTFIHTIVKLTLN
 VIENEQMENTQRAEHEEGQVKDLLAESRLRYKDYIQILSSPNFSLDQYCEQMWHREKRRQRN
 KGGPKWKHMQEMKKKRNRHRDLDELPRAVAT

Important features of the protein:**Signal peptide:**

amino acids 1-36

N-glycosylation sites.

amino acids 139-142, 607-610, 724-727

Tyrosine kinase phosphorylation site.

amino acids 571-576

Gram-positive cocci surface proteins 'anchoring' hexapeptide.

amino acids 32-37

CCCTGACCTCCCTGAGCCACACTGAGCTGGAAGCCGCAGAGGTCATCCTGGAGCATGCCACCGCCGGGGAGCAGA
CAACCTCCCAGGTAAGCTGGGAGCAAGACCTGAAGCTGTTTCTTCAGGAGCCTGGTGTATTTTCCCCACCCCAC
CTCAGCAGTTTCAGCCAGCAGGGACTGATCAGGTGTGTGTCTCTGGAGTGGGGAGCAGAAGGCGTGGCTGGCAAGA
GTGGCCTGGAGAAAAGAGGTTTCAGCGCTTGACCAGCCGAGCTGCCCGTGACTIONACAAGATCCAGAACCATGGGCATC
GGCTGAGGTGGGGGGGCACAGGTGTCTATGTGCACCTTCTTGTCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG
AGCCATTGAGGTTGCTATGGAGCTACAGAGGGGAGGAAAGGTATTTAAGGTAACAGTGTGGCACAATAGTTAA
GAGCACAGTTTTTTGGAGCTAGACCGACATAGGTTCAAATTTCTTCTGTTGCTTCTAGTTCTGTAGCCCCAGGT
AAGGGAGTGACTTAACCTCTCTGGACTTCAATTTCTCTCATCTACTAAAGTAGGGCCAATAATAGCACCCACCTCAT
AGGGAAGATTAAATGACATAATGTATGTGATGCACTAGCAAAAGTACCAGTCCCATAGTAAGTCATGCCCCACAC
TATTTCCACCCAGCCCTGTTTCTCTGCCTTCCCAACCAGGTACTGCAACGACTGGAGCAGAGGCGGCAGCAGGCTT
CAGAGCGGGAGGCTCCAAGCATAGAAACAGAGGTTACAGGAAGTGCAGAGAGCATCCGCCGGGCACAGGTGAGCC
AGGTGAAGGGGGCTGCCCGCTGGCCCTGCTGTCAGGGGGCTGGCTTAGATGTGGAGCGCTGGCTGAAGCCAGCCA
TGACCCAGGCCCAGGATGAGGTGGAGCAGGAGCGGGCGGTCACTGAGGTCGGTGTCCCAGAGGGACCTCTCTC
CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGGAATGTGAGGAGACGGGAGAGCTCTTTGAGGAGCCTGCC
CCCAAGCCCTGGCCACGAGGGGCCCTCCCCTGCCCTGCACACGTGGTATTTGCTATCAGGCAGGGCGTGAGGATG
AGCTGACAATCACGGAGGGTGAGTGGCTGGAGGTATAGAGGAGGGAGATGCTGACGAATGGGTCAAGGCTCGGA
ACCAGCAGCGGCAGGTTAGGCTTTGTCCCTGAGCGATATCTCAACTTCCCGGACCTCTCCCTCCCAGAGAGCAGCC
AAGACAGTGACAATCCCTGCGGGGACGAGCCACAGCATCTCTGGCACAGGCCCTGTACAGCTACACCGGACAGA
GTGCAGAGGAGCTGAGCTTCCCTGAGGGGGCACTCATCCGTCTGCTGCCCGGGCCCAAGATGAGTAGATGACG
GCTTCTGGAGGGGAGAATTTGGGGGCCGTGTTGGGGTCTTCCCCTCCCTGCTGGTGGAAGAGCTGCTTGGCCCC
CAGGGCCACCTGAACCTCTCTGACCCTGAACAGATGCTGCCGTCCCCTTCTCCTCCCAGCTTCTCCCCACCTGCAC
CTACCTTGTGTTGGATGGGCCCCCTGCACCTGTCTGCTGGGGACAAAGCCCTGGACTTCCCTGGGTTCTTG
ACATGATGGCACCTCGACTCAGGCCGATGCGTCCACACCTCCCCCGCGGCTAAAGCCCCGGATCCTGGCCACC
CAGATCCCCTCACCTGAGGCCAGGGAAGCCTTGACCCCCAGTGATGCTGCTCTCTTCAAGCTGTCTCAGA
CCACACCATCAATGATTCAGAGCAACACAGCCAAAAGCTGGAATCGCCCTTATTTCCACCTCACCTCAAGGGT
GGAAACTTGCCCCCTTCCCATTTCTAGAGCTGGAACCCACTCCTTTTTTTTCCCATTTGTTCTATCATCTCTAGGACC
GGAATACTACCTTCTTCTGTCTATGACCCTATCTAGGGTGGTGAAATGCCTGAAATCTCTGGGGCTGGAAACC
ATCCATCAAGGCTCTAGTAGTTCTGGCCACCTCTTTCCCACCTTGCTCCATGACCCACCCCACTCTGGATG
CCAGGTCAGCTGGGTTGGGCTGGGGAGGGAAGCAGGCTTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG
CTGTAATGGTCTGAGCGGATTTATTGACAATGAATAAAGGACAGGAGCCAGGCGAGGCTGGGCTCTTGTG
CTAAGAGGGCAGGGGGCCTACGGTGCTATTGCTTTAGGGGGCCACACCGGCAGGGGCTGCTCCCAGCTGCCAC
GCTCTATCATATGGAGCGAGGTGTTGGGGAAGGCGGGGCAGGCAGCCTGTTGCAGGCAGGGGAAGGAGAAGAGAC
TGAGGGGGCTGTGACCTCTCCTGAGGCCCCACAGCTGAGACTGTGCAACTCCAGGTGGAAGTAGAGCTGGTCCCTC
AGCTGGGGGCAAGTGCTGTCCAGTGGAGGGGAGGGCTTTCAGCCCCACCCACCCCTGGCCCTGCCAGCTGGTAG
TCCATCAGGCACAATGAAGGAGACTTGGAGAAGAGGAAGAATAACACTGTGCTTCTGTTCAAGCTGTGTCCAGC
TTTTCCCCTGGGGCTCCAGGACCTTCCCTACCTCCACCACCAACCAAGGATTTATAGCAAAGGCTAAGCCTGC
AGTTTTACTCTGGGGGTTTCAAGGAGCCGAAAGGCTTAAATAGTTTAAAGTAGGTATGGGAAGATGAGATTACCTCA
TTTAGGGCTCAGGCAGACTCACCTCACATACTCCCTGCTCCCTGTGGTAGAGACACCTGAGAGAAAGGGGAGGGG
TCAACAATGAGAGACCCAGGAGTAGTCTATCAGTGCCCCCAGAGTAGAGAGCAATAAGAGCCCAGCCCAGTGC
AGTCCCGGCTGTGTTTTCTACCTGGTGATCAGAAGTGCTGTTGCTTGGCTGGCTGCCATTTGCTCTTGTAGTGG
GCAGCCCTGGGCTTGGGCCCCCTCCCTCGGCCCTCAGTGTGGCTGTGCAAGCTCTGGGTTCCCTTCAAGTG
CACGAGGGGTTAGGCTGCTGTCCCTGAGTCTCCATTCTGTACTGGGGGGCTGGCTAGGACCTGGGCTGTGGCC
TCTCAGGGGGCAGCCTCTCCATGGCAGGCATCCCTGCCTTGGGCTGCCCTCCCCAGACCCCTGACCACCCCTG
GGTCTGTCCCCCACCAGAGCCCCAGCTCCTGTCTGTGGGGGAGCCATCACGGTGTTCGTGCAGTCCATAGCGCT
TCTCAATGTGTGTACCCCGGAACCTGGGAGGGGAGGGAACACTGGGGTTTAGGACCACAACCTCAGAGGCTGCTTG
GCCCTCCCCTGACACAGGACATCCTGAGTTTGGTGGCTACTTCCCTCTGGCCTAAGGTAGGGGAGGCCTTCTC
AGATTGTGGGACCATTTGTGTAGCTGACTTCTGTGTGAGCTCCCAGTCAGGAGGAAGAGAGCCAAAGGCCACTT
TTGGGATCAGGTGCCTGATCACTGGGCCCCCTACCTCAGCCCCCCTTTCCCTGGAGCACTGCCACCTGCCCA
CAGAGAACACAGTGGTCTCCCTGTCCGGGGGCGGCTTTTTCTTCTTGGAGCGTCCCTGACGGACAAGTGGAG
GCCTCTTGTGCGGCTGCAATGGATGCAAGGGGCTGCAGAGCCCAGGTGCACTGTGTGATGATGGGAGGGGGCTC
CGTCTGCAAGGCTGGAGGTGGCATCCACACTGGACAGCAGGAGGAGGGGAGTGAGGGTAACATTTCCATTTCCCT
TCATGTTTTGTTTCTTACGTTCTTTCAGACTGCTCCTTAAACCCCAAGCCCCAATTTCCCAAGCCCCATTT
TTTCTTGTCTTTATCTAATAAACTCAATATTAAG

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401
><subunit 1 of 1, 370 aa, 1 stop
><MW: 40685, pI: 4.53, NX(S/T): 0
MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLQRRQQASEREAPSIEQRLQEVRESIRRA
QVSQVKGAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELS
DFEECEETGELFEEPAPQALATRALPCPAHVVFYRYQAGREDELITITEGEWLEVIEEGDADEW
VKARNQHGEVGFVPERYLNFPDLSLPESSQSDNPGCAEPTAFLAQALYSYTGQSAEELSFP
EGALIRLLPRAQDGVDDGFWRGEFGGRVGVFPSLLVEELLGPPGPPPELSDPEQMLPSPSPPS
FSPPAPTSLVDGPPAPVLPDGDKALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT
```


[illegible]

FIGURE 180

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492

><subunit 1 of 1, 837 aa, 1 stop

><MW: 90167, pI: 8.39, NX(S/T): 1

MSQTGSHPGRGLAGRWLWGAQPCLLLPVPLSWLVWLLLLLLASLLPSARLASPLPREEEIV
 FPEKLNGSVLPGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGQAPELLGGAEP
 GTYLTGTINGDPESVASLHWDGGALLGVLYRGAELHLQPLEGGTPNSAGGPGAHLRRKSP
 ASGQGPMCNVKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHGAGLKRYLLTVMAA
 AAKAFKHPSIRNPVSLVVTRLVILGSGEEGPQVGPSAAQTLRSFCAWQRGLNTPEDSGPDHF
 DTAILFTRQDLCGVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHD
 NSKPCISLNGPLSTSRHVMAPVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHL
 PVTFPKDYDADRQCQLTFGPDSRHCPQLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCG
 PAQACMGGRCLHMDQLQDFNIPQAGGWGPWGPWGDCSRTC GGGVQFSSRDCTRVPVRNGGKY
 CEGRRTRFRSCNTEDCPTGSALTFREEQCAAYNHRTDLFKSFPGPMDWVPRYTGVAPQDQCK
 LTCQARALGYYYVLEPRVVDGTPCSPDSSSVCVQGRCIHAGCDRIIGSKKKFDKCMVCGGDG
 SGCSKQSGSFRKFRYGYNNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTL
 MPSPTDVVLPGAVSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPT
 PSTPRPTPDWLHRRRAQILEILRRRPWAGRK

Important features of the protein:

Signal peptide:

amino acids 1-48

N-glycosylation site.

amino acids 68-71

Glycosaminoglycan attachment site

amino acids 188-191, 772-775

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 182-185

Tyrosine kinase phosphorylation site.

amino acids 730-736

N-myristoylation sites.

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-
 172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-
 582, 679-684, 682-687, 763-768

Amidation sites.

amino acids 560-563, 834-837

Leucine zipper pattern.

amino acids 17-38, 24-45

Neutral zinc metalloproteinases, zinc-binding region signature.

amino acids 358-367

FIGURE 181

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGG
CAAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCC
AAGAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTTTGGTATCCTGGCCCTAACTCT
AATTGTCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAAGCCTATGACA
TGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTG
ACCAGAACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTT
TAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGA
TTAAAGTGATTCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACC
ACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGA
TTTTCTTAAAAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATC
CCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTT
CCTGCCAACGAAAAAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGT
AGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTG
AAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGT
CGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATA
CTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTGTAAGTGGTGGGTGGCCC
GCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAACTGCTGGCAACATAT
AATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCT
CTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACC
AAAAAAAAAAAAAAAAAAAA

FIGURE 182

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73727
><subunit 1 of 1, 317 aa, 1 stop
><MW: 37130, pI: 5.18, NX(S/T): 3
MAKNPPENCEDCHILNAEAFKSKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAY
DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKT
QIKVIPEFSEPEEEIDENEEITTTFFEQSVIWWVPAEKPIENRDFLKNKILEICDNVTMYWI
NPTLISVSELQDFEEEGEDLHFPA NEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDY
TENGIEFDPMLDERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWV
ARMLGRV
```

Important features of the protein:**Signal peptide:**

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 183

GCGGAACTGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCC
 CGGCTGGCCTAGGCAGGCAGCCGCACC**ATG**GCCAGCACGGCCGTGCAGCTTCTGGGCTTCCT
 GCTCAGCTTCCTGGGCATGGTGGGCACGTTGATCACCACCATCCTGCCGCACTGGCGGAGGA
 CAGCGCACGTGGGCACCAACATCCTCACGGCCGTGTCCTACCTGAAAGGGCTCTGGATGGAG
 TGTGTGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCC
 CCAAGACCTCCAGGCTGCCCGCGCCCTCATGGTCATCTCCTGCCTGCTCTCGGGCATAGCCT
 GCGCCTGCGCCGTTCATCGGGATGAAGTGACGCGCTGCGCCAAGGGCACACCCGCCAAGACC
 ACCTTTGCCATCCTCGGCGGCACCCTCTTCATCCTGGCCGGCCTCCTGTGCATGGTGGCCGT
 CTCCTGGACCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCAGCGGCATGA
 AGTTTGAGATTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGT
 GGCACCCTGCTTTGCCTGTCTTGCCAGGACGAGGCACCCTACAGGCCCTACCAGGCCCGGCC
 CAGGGCCACCACGACCACTGCAAACACCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAG
 ACAATCGGGCCCCCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTG
TGAGTCCCCACAGCCTGCTTCTCCCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTC
 AATGGAGGCAGGGGTTCAGCACAAAGTTTACTTCTGGGCAATTTTTGTATCCAAGGAAATA
 ATGTGAATGCGAGGAAATGTCTTTAGAGCACAGGGACAGAGGGGGAAATAAGAGGAGGAGAA
 AGCTCTCTATAACAAAGACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATAT
 TTATGTGGGTGATTTGATAACAAGTTTAATATAAAGTGACTTGGGAGTTTGGTCAGTGGGGT
 TGGTTTGTGATCCAGGAATAAACCTTGCGGATGTGGCTGTTTATGAAAAAAAAAAAAA

FIGURE 184

MASTAVQLLGFLLSFLGMVGTLITTLPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIY
QCQIYRSLLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAKGTPAKTTFAILGGTL
FILAGLLCMVAVSWTTNDVVQNFYNPLLP SGMKFEIGQALYLGFISSSLSLIGGTLLCLSCQ
DEAPYRPYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-103, 115-141, 160-182

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 185

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCG
 CAGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTCCT
 CCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCC
 AGGACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGC
 GTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGC
 CATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCC
 TGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCC
 AACATGACACTGACCTCCGGGATCATGTTTATTGTCTCAGGTCTTTGTGCAATTGCTGGAGT
 GTCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCG
 GCATGGGTGGGATGGTGCAGACTGTTTCCAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGTG
 GGCTGGGTGCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGG
 CCTGGCACCAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTG
 CCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAG
 ATATACGATGGAGGTGCCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTA
 TGTG**TAA**TGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAA
 AACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGC
 CTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCC
 ACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTAT
 TTCTTTTTTTAAATATACTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCAC
 ATTTTGATGATTTAGACAGACTCCCCCTCTTCTCCTAGTCAATAAACCCATTGATGATCTA
 TTTCCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTT
 CTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAA
 GAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTG
 TGATCTTAAAGTTACCAAACCAAAGTCATTTTTCAGTTTGAGGCAACCAAACCTTTCTACTG
 CTGTTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAG
 TCCTCTTTCTGTGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAAT
 TTAAGTCCTAAATATAGTTAAATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGA
 AATAGCCTCACCCCTACATGTGGATAGAAGGAAATGAAAAATAATTGCTTTGACATTGTCT
 ATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAAAGCTCACACCTGTAATC
 CTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTG
 GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGGTG
 GCATACACCTGTAGTCCCAGCATTCGCGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGG
 AGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGA
 TCCTGTCTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAA
 ACTAATTCTTTAA

100
 90
 80
 70
 60
 50
 40
 30
 20
 10
 0


```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734
><subunit 1 of 1, 261 aa, 1 stop
><MW: 27856, pI: 8.50, NX(S/T): 1
MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGWLWRS CVRQSSGFTE
CRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIM
FIVSGLCAIAGVS VFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTL
IGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTE
DEVQSYPSKHDYV
```

amino acids 1-23

amino acids 81-100, 121-141, 173-194

FIGURE 187

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCG
GAGTCCAGCTGGCTAAAACATCCCAGAGGATA**ATGG**CAACCCATGCCTTAGAAATCGCTG
GGCTGTTTCTTGGTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGG
AGAGTGTGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAGAACTTCTGGGAAGGACTGTG
GATGAATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCTGCTGG
CTCTTTCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTC
TTGGCTTTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAA
GGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGC
TCATCCCTGTGAGCTGGGTGGCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAAT
GTTGCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGTCT
GATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACA
GATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCG
AGCGTCTACTCCAGAAGTCAGTATGT**GTAG**TTGTGTATGTTTTTTAACTTTACTATAAAGC
CATGCAAAATGACAAAATCTATATTACTTTCTCAAATGGACCCCAAAGAACTTTGATTTA
CTGTTCTTAACTGCCTAATCTTAATTACAGGAACGTGTCATCAGCTATTTATGATTCTATAA
GCTATTTTACAGCAGAATGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAAT
TTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTATCATTTACTTCAAATGACATTGCT
AAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAG
TGTAACATTTATATCTCACATAGAGACATGCTTATATGGTTTTATTTAAATGAAATGCCAG
TCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTG
AAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGA
AGATTAATAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTG
TTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTT
TTCTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCA
AACTGCTTTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATG
GTTTTAGGAAAGTGAAAATATTTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAA
GAAATCATATATGTATGGATATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATC
AATATAAATAAAAGAGCAGAAAAATATGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACA
ACAAAAAAGTTGTCCTTTGAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATTG
TCATTTTTGTTCTGTGAAAAATAAATTTCTTCTGTACCATTCTGTTTAGTTTTACTAAA
ATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAACTGACAATCCAATTTGA
AAGTTTGTGTGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTTTATACATTTATA
TTAATAAATTGTACATTTTTCTAATT

FIGURE 188

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRM
QCKIYDSLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGI
IFIITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVF
CCNEKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

amino acids 82-101, 118-145, 164-188

FIGURE 189

TCGCC**ATG**GCCTCTGCCGGAATGCAGATCCTGGGAGTCGTCTTGACACTGCTGGGGCTGGGTG
AATGGCCTGGTCTCCTGTGCCCTGCCCATGTGGAAGGTGACCGCTTTCATCGGCAACAGCAT
CGTGGTGGCCCAGGTGGTGTGGGAGGGCCTGTGGATGTCCTGCGTGGTGCAGAGCACCGGCC
AGATGCAGTGCAAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGGCTGCACGT
GCCCTCTGTGTATCGCCCTCCTTGTGGCCCTGTTTCGGCTTGCTGGTCTACCTTGCTGGGGC
CAAGTGTACCACCTGTGTGGAGGAGAAGGATTCCAAGGCCCGCCTGGTGCTCACCTCTGGGA
TTGTCTTTGTATCTCAGGGGTCTGACGCTAATCCCCGTGTGCTGGACGGCGCATGCCATC
ATCCGGGACTTCTATAACCCCTGGTGGCTGAGGCCCAAAGCGGGAGCTGGGGGCCTCCCT
CTACTTGGGCTGGGCGGCCTCAGGCCTTTTGTGCTGGGTGGGGGGTTGCTGTGCTGCACTT
GCCCCTCGGGGGGGTCCCAGGGCCCCAGCCATTACATGGCCCGCTACTCAACATCTGCCCCT
GCCATCTCTCGGGGGCCCTCTGAGTACCCTACCAAGAATTACGTC**TGA**CGTGGAGGGGAATG
GGGGCTCCGCTGGCGCTAGAGCCATCCAGAAGTGGCAGTGCCCAACAGCTTTGGGATGGGTT
CGTACCTTTTGTCTCTGCCTCCTGCTATTTTTCTTTGACTGAGGATATTTAAAATTCAATTT
GAAAACTGAGCCAAGGTGTTGACTCAGACTCTCACTTAGGCTCTGCTGTCTCTCACCCCTTGG
ATGATGGAGCCAAAGAGGGGATGCTTTGAGATTCTGGATCTTGACATGCCATCTTAGAAGC
CAGTCAAGCTATGGAATAATGCGGAGGCTGCTTGCTGTGCTGGCTTTGCAACAAGACAGAC
TGTCCCCAAGAGTTCTGCTGCTGCTGGGGGCTGGGCTTCCCTAGATGTCAGTGGACAGCTG
CCCCCATCCTACTCAGGTCTCTGGAGCTCCTCTCTTACCCCTGGAAAAACAAATCATCTG
TTAACAAAGGACTGCCACCTCCGGAACCTTCTGACCTCTGTTTCCCTCCGTCCTGATAAGACG
TCCACCCCCCAGGGCCAGGTCCCAGCTATGTAGACCCCCGCCCCACCTCCAACACTGCACC
CTTCTGCCCTGCCCCCTCGTCTCACCCCTTTACACTCACATTTTTATCAAATAAAGCATG
TTTTGTTAGTGCA

FIGURE 190

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73736

><subunit 1 of 1, 220 aa, 1 stop

><MW: 23292, pI: 8.43, NX(S/T): 0

MASAGMQILGVVLTLLGWVNGLVSCALPMWKVTAFIGNSIVVAQVVWEGLWMSCVVQSTGQM
QCKVYDSSLALPQDLQAARALCVIALLLVALFGLLVYLAGAKCTTCVEEKDSKARLVLTSGIV
FVISGVLTLIPVCWTAHAIIRDFYNPLVAEAQKRELGASLYLGWAASGLLLLGGGLLCCTCP
SGGSQGP SHYMARYSTSAPAISRG PSEYPTKNYV

Transmembrane domains:

amino acids 8-30 (type II), 82-102, 121-140, 166-186

1001393 63631001

FIGURE 191

GCCTAAGGAGAACATCATCAAAGACTTCTCTAGACTCAAAAGGCTTCCACGTTCTACATCTTG
AGCATCTTCTACCACTCCGAATTGAACCAGTCTTCAAAGTAAAGGCAATGGCATTTTATCCC
TTGCAAATTGCTGGGCTGGTTCTTGGGTTCCTTGGCATGGTGGGGACTCTTGCCACAACCCCT
TCTGCCTCAGTGGTGGAGTATCAGCTTTTGTGGCAGCAACATTATTGTCTTTGAGAGGCTC
TGGGAAGGGCTCTGGATGAATTGCATCCGACAAGCCAGGGTCCGGTTGCAATGCAAGTTCTA
TAGCTCCTTGTTGGCTCTCCCGCCTGCCCTGGAAACAGCCCCGGGCCCTCATGTGTGTGGCTG
TTGCTCTCTCCTTGATCGCCCTGCTTATTGGCATCTGTGGCATGAAGCAGGTCCAGTGCACA
GGCTCTAACGAGAGGGGCCAAAGCATACCTTCTGGGAACCTCAGGAGTCTCTTCATCCTGAC
GGGTATCTTCGTTCTGATTCCGGTGAGCTGGACAGCCAATATAATCATCAGAGATTTCTACA
ACCCAGCCATCCACATAGGTGAGAAACGAGAGCTGGGAGCAGCACTTTTCCTTGGCTGGGCA
AGCGCTGCTGTCTCTTCATTGGAGGGGGTCTGCTTTGTGGATTTTGCTGCTGCAACAGAAA
GAAGCAAGGGTACAGATATCCAGTGCCTGGCTACCGTGTGCCACACACAGATAAGCGAAGAA
ATACGACAATGCTTAGTAAGACCTCCACCAGTTATGTCTTAATGCCTCCTTTTGGCTCCAAGT
ATGGACTATGGTCAATGTTTTTTATAAAGTCCTGCTAGAACTGTAAGTATGTGAGGCAGGA
GAACTTGCTTTATGTCTAGATTTACATTGATACGAAAGTTTCAATTTGTTACTGGTGGTAGG
AATGAAAATGACTTACTTGGACATTCTGACTTCAGGTGTATTAAATGCATTGACTATTGTTG
GACCCAATCGCTGCTCCAATTTTCATATTCTAAATTCAAGTATACCCATAATCATTAGCAAG
TGTACAATGATGGACTACTTATTACTTTTTGACCATCATGTATTATCTGATAAGAATCTAAA
GTTGAAATTGATATTCTATAACAATAAAACATATACCTATTCTA

FIGURE 192

MNCIRQARVRLQCKFYSSLLALPPALETARALMCVAVALSLIALLIGICGMKQVQCTGSNER
AKAYLLGTSGVLFILTGIFVLIPVSWTANIIIRDFYNPAIHIGQKRELGAALFLGWASAAVL
FIGGGLLCGFCCCNRKKQGYRYPVPGYRVPHTDKRRNTTMLSKTSTSYV

amino acids 161-164

FIGURE 193

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGCATGAAGATCACTGGGGGT
CTCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCC
AAAAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCA
CATACTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACTTGTGTACC
GAGAGCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGCTAAATTCTCCA
TGGACATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTG
AGTTTCTTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTGAGAGTAATCTTG
ACTGAATGGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCAT
TTTTTTTTTAACACGTCAATAAAAAAATAATCTCCCAGA

10123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

FIGURE 194

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73739
```

```
><subunit 1 of 1, 85 aa, 1 stop
```

><MW: 9232, pI: 7.94, NX(S/T): 0

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGN
 ECHLCTESLKSNGRVQFLHDGSC

Signal peptide:

amino acids 1-19

FIGURE 195

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCCCTCGCCCTCCTGCTCGCGCCC
CGCCGCC**ATG**GCTGCCTCCCCGCGCGGCCTGCTGTCCTGGCCCTGACCGGGCTGGCGCTGC
TCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAAATAAACTCAAGCTGATGCTTCAA
AAACGAGAAGCACCTGTTCCAATAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGA
ATTCCTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCCGAGGTGC
AGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATC
ACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCA
CTATGATGAAGACTCTGCAATTGGTCCCCGGAGCCCCCTACGGCTTTAGGCATGGAGCCAGCG
TCAACTACGATGACTAC**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATT
TCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTACGCAGAT
CTTTTCTACCTACTTTGTGTGATCAAAAAAGAAGAGTTAAAACAACACATGTAAATGCCTTT
TGATATTTTCATGGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTAAAAAGA

1004333104

FIGURE 196

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742
```

><subunit 1 of 1, 148 aa, 1 stop

><MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLLCWGPGGISGNKCLKMLQKREAPVPTKTKVAVDENKAKEFL
GSLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRNGHEYGYGDIYQRHYD
EDSAIGPRSPYGFRHGASVNYDDY

Signal peptide:

amino acids 1-30

FIGURE 197

CGGCTCGAGCCCGCCCGGAAGTGCCCGAGGGGCGCGATGGAGCTGGGGGAGCCGGGCGCTC
 GGTAGCGCGGGCGGGCAAGGCAGGCGCC**ATG**ACCCTGATTGAAGGGGTGGGTGATGAGGTGAC
 CGTCCTTTTCTCGGTGCTTGCCTGCCTTCTGGTGCTGGCCCTTGCCTGGGTCTCAACGCACA
 CCGCTGAGGGCGGGGACCCACTGCCCCAGCCGTAGGGACCCCAACGCCATCCCAGCCCAGC
 GCAGCCATGGCAGCTACCGACAGCATGAGAGGGGAGGCCCCAGGGGCAGAGACCCCCAGCCT
 GAGACACAGAGGTCAAGCTGCACAGCCAGAGCCCAGCACGGGGTTACAGCAACACCGCCAG
 CCCC GGACTCCCCGCAGGAGCCCCTCGTGCTACGGCTGAAATTCCTCAATGATTGAGAGCAG
 GTGGCCAGGGCCTGGCCCCACGACACCATTGGCTCCTTGAAAAGGACCCAGTTTCCCGGCCG
 GGAACAGCAGGTGCGACTCATCTACCAAGGGCAGCTGCTAGGCGACGACACCCAGACCCTGG
 GCAGCCTTCACCTCCCTCCCAACTGCGTTTCTCCACTGCCACGTGTCCACGAGAGTCGGTCCC
 CCAAATCCCCCCTGCCCCGCCGGGGTCCGAGCCCGGCCCTCCGGGCTGGAAATCGGCAGCCT
 GCTGCTGCCCCCTGCTGCTCCTGCTGTTGCTGCTGCTCTGGTACTGCCAGATCCAGTACCGGC
 CCTTCTTTCCCCTGACCGCCACTCTGGGCCTGGCCGGCTTCACCCTGCTCCTCAGTCTCCTG
 GCCTTTGCCATGTACCGCCCG**TAG**TGCCTCCGCGGGCGCTTGGCAGCGTCGCCGGCCCCCTCC
 GGACCTTGCTCCCCGCGCCGCGGCGGGAGCTGCTGCCTGCCCAGGCCCGCCTCTCCGGCCTG
 CCTCTTCCCGCTGCCCTGGAGCCCAGCCCTGCGCCGCAGAGGACTCCCGGGACTGGCGGAGG
 CCCC GCCCTGCGACCGCCGGGGCTCGGGGCCACCTCCCGGGGCTGCTGAACCTCAGCCCGCA
 CTGGGAGTGGGCTCCTCGGGGTGCGGCATCTGCTGTGCTGCCTCGGCCCCGGGCAGAGCCG
 GGCCGCCCCGGGGGCCCGTCTTAGTGTTCTGCCGGAGGACCCAGCCGCTCCAATCCCTGAC
 AGCTCCTTGGGCTGAGTTGGGGACGCCAGGTGCGTGGGAGGCTGGTGAAGGGGAGCGGGGAG
 GGGCAGAGGAGTTCCCCGGAACCCGTGCAGATTAAAGTAACTGTGAAGTTTAAAAAAAAAA
 AAAAAAA

FIGURE 198

MTLIEGVGDEVTVLFSVLACLLVLALAWVSTHTAEGGDPLPQPSGTPTPSQPSAAMAATDSM
RGEAPGAETPSLRHRGQAAQPEPSTGFTATPPAPDSPQEPLVRLKFLNDSEQVARAWPHDT
IGSLKRTQFPGREQQVRLIYQGQLLGDDTQTLGSLHLPNCVLHCHVSTRVGPPNPPCPPGS
EPGPSGLEIGSLLLPLLLLLLLLLLWYCQIQYRPFPLTATLGLAGFTLLLSLLAFAMYRP

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 195-217

1001334404

FIGURE 199

GAGATTGGAAACAGCCAGGTTGGAGCAGTGAGTGAGTAAGGAAACCTGGCTGCCCTCTCCAG
ATTCCCCAGGCTCTCAGAGAAGATCAGCAGAAAGTCTGCAAGACCCTAAGAACCATCAGCCC
TCAGCTGCACCTCCTCCCCCTCCAAGGATGACAAAGGCGCTACTCATCTATTTGGTCAGCAGC
TTTCTTGCCCTAAATCAGGCCAGCCTCATCAGTCGCTGTGACTTGGCCCAGGTGCTGCAGCT
GGAGGACTTGGATGGGTTTGAGGGTTACTCCCTGAGTGACTGGCTGTGCCTGGCTTTTGTGG
AAAGCAAGTTCAACATATCAAAGATAAATGAAAATGCGGATGGAAGCTTTGACTATGGCCTC
TTCCAGATCAACAGCCACTACTGGTGCAACGATTATAAGAGTTACTCGGAAAACCTTTGCCA
CGTAGACTGTCAAGATCTGCTGAATCCCAACCTTCTTGCAGGCATCCACTGCGCAAAAAGGA
TTGTGTCCGGAGCACGGGGGATGAACAACCTGGGTAGAATGGAGGTTGCACTGTTTCAAGGCCGG
CCACTCTCCTACTGGCTGACAGGATGCCGCCTGAGATGAAACAGGGTGCGGGTGCACCGTGG
AGTCATTCCAAGACTCCTGTCCTCACTCAGGGATTCTTCATTTCTTCTTCTCCTACTGCCTCCA
CTTCATGTTATTTTCTTCCCTTCCCATTACAACTAAAACCTGACCAGAGCCCCAGGAATAAA
TGGTTTTCTTGGCTTCCTCCTTACTCCCATCTGGACCCAGTCCCCTGGTTCCTGTCTGTTAT
TTGTAAACTGAGGACCACAATAAAGAAATCTTTATATTTATCG

FIGURE 200

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746

><subunit 1 of 1, 148 aa, 1 stop

><MW: 16896, pI: 6.05, NX(S/T): 1

MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEGYSLSDWLCCLAFVESKFENISKI
NENADGSFDYGLFQINSHYWCNDYKSYSENLCNVDCQDLLNPNNLAGIHCAKRIVSGARGMN
NWVEWRLHCSGRPLSYWLTGCRLR

Signal peptide:

amino acids 1-18

TCTGACCTGACTGGAAGCGTCCAAAGAGGGGACGGCTGTACGCCCTGCTTGACTGAGAACCCCA
 CCAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGGAAAGAAACACCTGAG
 CAGAATGGAATCATTATTTTTTTTCCCAAGGAGAAAACCGGGGTAAAGGGAGGGAAGCAATTC
 AATTTGAAGTCCCTGTGAATGGGCTTTCAGAAGGCAATTAAGAAATCCACTCAGAGAGGAC
 TTGGGGTGAAACTTGGGTCTGTGGTTTTCTGATTGTAAGTGGAAGCAGGTCTTGACACACGC
 TGTGGCAAATGTCAGGACCAGGTTAAGTGAAGTGGCAGAAAACTTCCAGGTGGAACAAGCA
 ACCCATGTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGAGAAAGCTAACTTGAACATGAC
 CTGTTGCATTTGGCAAGTTCTAGCAAC**ATG**CTCCTAAGGAAGCGATACAGGCACAGACCATG
 CAGACTCCAGTTTCTCCTGCTGCTCCTGATGCTGGGATGCGTCCTGATGATGGTGGCGATGT
 TGCACCCTCCCCACCACACCCTGCACCAGACTGTCACAGCCCAAGCCAGCAAGCACAGCCCT
 GAAGCCAGGTACCGCCTGGACTTTGGGGAAATCCAGGATTGGGTACTGGAAGCTGAGGATGA
 GGGTGAAGAGTACAGCCCTCTGGAGGGCCTGCCACCCTTTATCTCACTGCGGGAGGATCAGC
 TGCTGGTGGCCGTGGCCTTACCCAGGCCAGAAGGAACCAGAGCCAGGGCAGGAGAGGTGGG
 AGCTACCGCCTCATCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGACTGGGG
 GGCTGATGAGGACGGGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTCAGCCTGGACCCAC
 GTGGCCTCCAGGAGGCACTCAGTGCCCGCATCCCCCTCCAGAGGGCTCTGCCCGAGGTGCGG
 CACCCACTGTGTCTGCAGCAGCACCCCTCAGGACAGCCTGCCACAGCCAGCGTCATCCTCTG
 TTTCCATGATGAGGCCTGGTCCACTCTCCTGCGGACTGTACACAGCATCCTCGACACAGTGC
 CCAGGGCCTTCCTGAAGGAGATCATCCTCGTGGACGACCTCAGCCAGCAAGGACAACCTCAAG
 TCTGCTCTCAGCGAATATGTGGCCAGGCTGGAGGGGGTGAAGTTACTCAGGAGCAACAAGAG
 GCTGGGTGCCATCAGGGCCCCGGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCT
 TCATGGATGCCCACTGCGAGTGCCACCCAGGCTGGCTGGAGCCCCCTCCTCAGCAGAATAGCT
 GGTGACAGGAGCCGAGTGGTATCTCCGGTGATAGATGTGATTGACTGGAAGACTTTCCAGTA
 TTACCCCTCAAAGGACCTGCAGCGTGGGGTGTTGGACTGGAAGCTGGATTTCCACTGGGAAC
 CTTTGCCAGAGCATGTGAGGAAGGCCCTCCAGTCCCCCATAAGCCCCATCAGGAGCCCTGTG
 GTGCCCGGAGAGGTGGTGGCCATGGACAGACATTACTTTCAAACACTGGAGCGTATGACTC
 TCTTATGTGCTGCGAGGTGGTGAACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTG
 GCTCTGTTGAAATCCTTCCCTGCTCTCGGGTAGGACACATCTACCAAATCAGGATTCCCAT
 TCCCCCCTCGACCAGGAGGCCACCCTGAGGAACAGGGTTTCGCATTGCTGAGACCTGGCTGGG
 GTCATTCAAAGAAACCTTCTACAAGCATAGCCCAGAGGCCTTCTCCTTGAGCAAGGCTGAGA
 AGCCAGACTGCATGGAACGCTTGCAGCTGCAAAGGAGACTGGGTGTGCGGACATTCCACTGG
 TTTCTGGCTAATGTCTACCCTGAGCTGTACCCATCTGAACCCAGGCCCAGTTTCTCTGGAAA
 GCTCCACAACACTGGACTTGGGCTCTGTGCAGACTGCCAGGCAGAAGGGGACATCCTGGGCT
 GTCCCATGGTGTGGCTCCTTGACGTGACAGCCGGCAGCAACAGTACCTGCAGCACACCAGC
 AGGAAGGAGATTCACTTTGGCAGCCCACAGCACCTGTGCTTTGCTGTCAGGCAGGAGCAGGT
 GATTCTTCAGAACTGCACGGAGGAAGGCCTGGCCATCCACCAGCAGCACTGGGACTTCCAGG
 AGAATGGGATGATTGTCCACATTCTTTCTGGGAAATGCATGGAAGCTGTGGTGCAAGAAAAC
 AATAAAGATTTGTACCTGCGTCCGTGTGATGGAAAAGCCCGCCAGCAGTGGCGATTTGACCA
 GATAAATGCTGTGGATGAACGAT**TGA**ATGTCAATGTCAGAAGGAAAAGAGAATTTTGGCCATC
 AAAATCCAGCTCCAAGTGAACGTAAAGAGCTTATATATTTTCATGAAGCTGATCCTTTTGTGT
 GTGTGCTCCTTGTGTTAGGAGAGAAAAAGCTCTATGAAAGAATATAGGAAGTTTCTCCTTT
 TCACACCTTATTTTCATTGACTGCTGGCTGCTTA

FIGURE 202

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760

><subunit 1 of 1, 639 aa, 1 stop

><MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFG
ESQDWVLEAEDEGEYSPLEGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPR
RQDKEAPKRWDGADEDEGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQOHP
QDSLPTASVILCFHDEAWSTLLRTVHSILDTVPRAFLKEIILVDDLSQQGQLKSALSEYVAR
LEGVKLLRSNKRLGAIRARMLGATRATGDVLVFMDAHCECHPGWLEPLLSRIAGDRSRVVSP
VIDVIDWKTFQYYPSKDLQRGVLDWKLDHFHWEPLPEHVRKALQSPISPIRSPVVPGEVVAMD
RHYFQNTGAYDSLMSLRGGENLELSFKAWLCGGSV EILPCSRVGHIYQNQDSHSPLDQEATL
RNRVRIAETWLGSFKETFYKHSPEAFSLSKA EKPDCMERLQLQRRLLGCRTFHWFLANVYPEL
YPSEPRPSFSGKLHNTGLGLCADCQAEGDILGCPMVLAPCSDSRQQQYLQHTSRKEIHFGSP
QHLCFAVRQEQVILQNCTEEGLAIHQHWD FQENGMIVHILSGKCMEAVVQENNKDLYLRPC
DGKARQQWRFDQINAVDER

Signal peptide:

amino acids 1-28

FIGURE 203

CGCCAAGCATGCAGTAAAGGCTGAAAATCTGGGTACAGCTGAGGAAGACCTCAGAC**ATGGA**
 GTCCAGGATGTGGCCTGCGCTGCTGCTGTCCACCTCCTCCCTCTCTGGCCACTGCTGTTGC
 TGCCCCCTCCCACCGCCTGCTCAGGGCTCTTCATCCTCCCCTCGAACCCCACCAGCCCCAGCC
 CGCCCCCGGTGTGCCAGGGGAGGCCCCCTCGGCCCCACGTATGTGTGCGTGTGGGAGCGAGC
 ACCTCCACCAAGCCGATCTCCTCGGGTCCCAAGATCACGTGCGCAAGTCCTGCCTGGCACTG
 CACCCCCAGCCACCCCATCAGGCTTTGAGGAGGGGGCCGCCCTCATCCCAATACCCCTGGGCT
 ATCGTGTGGGGTCCCACCGTGTCTCGAGAGGATGGAGGGGACCCCAACTCTGCCAATCCCGG
 ATTTCTGGACTATGGTTTTTGCAGCCCCTCATGGGCTCGCAACCCCACACCCCAACTCAGACT
 CCATGCGAGGTGATGGAGATGGGCTTATCCTTGGAGAGGCACCTGCCACCCTGCGGCCATTC
 CTGTTCCGGGGGCCGTGGGGAAGGTGTGGACCCCCAGCTCTATGTCACAATTACCATCTCCAT
 CATCATTGTTCTCGTGGCCACTGGCATCATCTTCAAGTTCTGCTGGGACCGCAGCCAGAAGC
 GACGCAGACCCTCAGGGCAGCAAGGTGCCCTGAGGCAGGAGGAGAGCCAGCAGCCACTGACA
 GACCTGTCCCCGGCTGGAGTCACTGTGCTGGGGGCCCTTCGGGGACTCACCTACCCCCACCCC
 TGACCATGAGGAGCCCCGAGGGGGACCCCGGCCTGGGATGCCCCACCCCAAGGGGGCTCCAG
 CCTTCCAGTTGAACCGGT**TGA**GGGCAGGGGCAATGGGATGGGAGGGCAAAGAGGGGAAGGCAAC
 TTAGGTCTTCAGAGCTGGGGTGGGGGTGCCCTCTGGATGGGTAGTGAGGAGGCAGGCGTGGC
 CTCCCACAGCCCCTGGCCCTCCCAAGGGGGCTGGACCAGCTCCTCTCTGGGAGGCACCCTTC
 CTTCTCCCAGTCTCTCAGGATCTGTGTCTATTCTCTGCTGCCCATAACTCCAACCTCTGCCC
 TCTTTGGTTTTTTCTCATGCCACCTTGTCTAAGACAACTCTGCCCTCTTAACCTTGATTCCC
 CCTCTTTGTCTTGAACCTTCCCCCTTCTATTCTGGCCTACCCCTTGGTTCCCTGACTGTGCCCTT
 TCCCTCTTCCTCTCAGGATTCCCCTGGTGAATCTGTGATGCCCCCAATGTTGGGGTGCAGCC
 AAGCAGGAGGCCAAGGGGGCCGGCACAGCCCCCATCCCACTGAGGGTGGGGCAGCTGTGGGGA
 GCTGGGGGCCACAGGGGCTCCTGGCTCCTGCCCTTGACACACCACCCGGAACACTCCCCAGCC
 CCACGGGCAATCCTATCTGCTCGCCCTCCTGCAGGTGGGGGCCCTCACATATCTGTGACTTCG
 GGTCCCTGTCCCCACCCTTGTGCACTCACATGAAAGCCTTGACACTCACCTCCACCTTCAC
 AGGCCATTTGCACACGCTCCTGCACCCCTCTCCCCGTCCATACCGCTCCGCTCAGCTGACTCT
 CATGTTCTCTCGTCTCACATTTGCACTCTCTCCTTCCCACATTCTGTGCTCAGCTCACTCAG
 TGGTCAGCGTTTTCCTGCACACTTTACCTCTCATGTGCGTTTCCCGGCCCTGATGTTGTGGTGG
 TGTGCGGCGTGCTCACTCTCTCCCTCATGAACACCCACCCACCTCGTTTCCGCAGCCCCCTGC
 GTGCTGCTCCAGAGGTGGGTGGGAGGTGAGCTGGGGGCTCCTTGGGGCCCTCATCGGTCATGG
 TCTCGTCCCATTCCACACCATTGTTTCTGTCTCCCCATCCTACTCCAAGGATGCCGGCA
 TCACCCTGAGGGCTCCCCCTTGGGAATGGGGTAGTGAGGCCCCAGACTTCACCCCCAGCCCA
 CTGCTAAAATCTGTTTTCTGACAGATGGGTTTTGGGGAGTCGCCTGCTGCACTACATGAGAA
 AGGGAATCCCATTTGCCCTTCCCTTTCTCCTACAGTCCCTTTTGTCTTGTCTGTCTGGCTG
 TCTGTGTGTGTGCCATTCTCTGGACTTCAGAGCCCCCTGAGCCAGTCCTCCCTTCCCAGCCT
 CCCTTTGGGCCTCCCTAACTCCACCTAGGCTGCCAGGGACCGGAGTCAGCTGGTTCAAGGCC
 ATCGGGAGCTCTGCCTCCAAGTCTACCCCTCCCTTCCCGGACTCCCTCCTGTCCCCTCCTTT
 CCTCCCTCCTTCCCTTCCACTCTCCTTCCCTTTGCTTCCCTGCCCTTTCCCCCTCCTCAGGTT
 CTTCCCTCCTTCTCACTGGTTTTTCCACCTTCCCTCCTTCCCTTCTTCCCTGGCTCCTAGGCT
 GTGATATATATTTTTGTATTATCTTTTCTTCTTCTTGTGGTGATCATCTTGAATTACTGTG
 GGATGTAAGTTTTCAAATTTTCAAATAAAGCCTTTGCAAGATAA

FIGURE 204

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393
><subunit 1 of 1, 243 aa, 1 stop
><MW: 26266, pI: 8.43, NX(S/T): 1
MRPQGPAASPQRLRGLLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPA
PGRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGPLPIEAIYYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 205

GTTAACCAGCGCAGTCCTCCGTGCGTCCCGCCCGCCGCTGCCCTCACTCCCGGCCAGG**ATGG**
CATCCTGTCTGGCCCTGCGCATGGCGCTGCTGCTGGTCTCCGGGGTTCTGGCCCTGCGGTG
CTCACAGACGATGTTCCACAGGAGCCCGTGCCACGCTGTGGAACGAGCCGGCCGAGCTGCC
GTCGGGAGAAGGCCCCGTGGAGAGCACCAGCCCCGGCCGGGAGCCCGTGGACACCGGTCCCC
CAGCCCCCACCCTCGCGCCAGGACCCGAGGACAGCACC GCGCAGGAGCGGCTGGACCAGGGC
GGCGGGTCGCTGGGGCCCGGCGCTATCGCGGCCATCGTGATCGCCGCCCTGCTGGCCACCTG
CGTGGTGCTGGCGCTCGTGGTCGTCGCGCTGAGAAAGTTTTCTGCCTCC**TGA**AGCGAATAAA
GGGGCCGCGCCCGGCCGCGGCGCGACTCGGCAAAAAAAAAAAAAAAAA

FIGURE 206

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398
><subunit 1 of 1, 121 aa, 1 stop
><MW: 12073, pI: 4.11, NX(S/T): 0
MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVESTSPGREPVDTG
PPAPTVPAGPEDSTAQERLDQGGGSLGPGAIAAIVIAALLATCVVLALVVVALRKFSAS
```

Important features of the protein:**Signal peptide:**

amino acids 1-19

Transmembrane domain:

amino acids 91-110

Glycosaminoglycan attachment site.

amino acids 44-47

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 116-119

N-myristoylation site.

amino acids 91-96

FIGURE 207

GGCCGTTGGTTGGTGCGCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGG
CGGGCCGGGACGGGCATGGCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGGCCCA
CGGCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGA
ACTTCAAGTCCTGGTGGGTGGGCGACATCCCCGTGTGAGGGGCGCTGCTCACCGACTGGAGC
GACGACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGAGAAGCTGGA
CCAAGTGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCC
CCGGGTATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAAC
GCCATCATCGAAAGGCACCTGGCACACAGGAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGG
ACCCAGCCTAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCCTAAGTAGCCC
CCAGAGGCGCTGGGAGTGTTGCCACCGCCCTCCCCTGAAGTTTGCTCCATCTCACGCTGGGG
GTCAACCTGGGGACCCCTTCCCTCCGGGCCATGGACACACATACATGAAAACCAGGCCGCAT
CGACTGTCAGCACCGCTGTGGCATCTTCCAGTACGAGACCATCTCCTGCAACAACCTGCACAG
ACTCGCACGTCGCTGCTTTGGCTATAACTGCGAGTAGGGCTCAGGCATCACACCCACCCGT
GCCAGGGCCCTACTGTCCCTGGGGTCCCAGGCTCTCCTTGGAGGGGGCTCCCCGCCTTCCAC
CTGGCTGTCATCGGGTAGGGCGGGGCCGTGGGTTGAGGGGCGCACCACTTCCAAGCCTGTGT
CCCACAGGTCTCGGCGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACTACATAAATAAC
TGGCACAAGTAAGTCCCCCTCCTCAAACCAACACAGGCAGTGTGTGTATGTGAGCACCTCGTG
GGTGAGTATGTGTGGGGCACAGGCTGGCTCCCTCAGCTCCACGTCTAGAGGGGGCTCCCGA
GGAGGTGGAACCTCAACCCAGCTCTGCGCAGGAGGCGGCTGCAGTCCTTTTCTCCCTCAAAG
GTCTCCGACCCCTCAGCTGGAGGCGGGCATCTTTCCTAAAGGGTCCCCATAGGGTCTGGTTCC
ACCCCATCCCAGGTCTGTGGTCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCC
ATGGAGGGGGCTGACTGCCCCACATTGCCTTTCAGACAGGACACGAGCATGAGGTAAGGCCGC
CCTGACCTGGACTTCAGGGGGAGGGGGTAAAGGGAGAGAGGAGGGGGGCTAGGGGGTCTCT
AGATCAGTGGGGGCACTGCAGGTGGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTCAC
CTCTGCAACCACACCCATGTGGTGGTTTCATGAACAGACCACGCTCCTCTGCCTTCTCCTGG
CCTGGGACACACAGAGCCACCCCGGCCTTGTGAGTGACCCAGAGAAGGGAGGCCCTCGGGAGA
AGGGGTGCTCGTAAGCCAACACCAGCGTGCCGCGGCCTGCACACCCTTCGGACATCCCAGGC
ACGAGGGTGTGCTGGATGTGGCCACACATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCCT
GGGGCCCCCAGGGAGGGAGGCAGGGGGTGGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCC
CCGCAGCCTGGTATCGCCAGCCTTAAGGTGTCTGGAGCCCCCACACTTGGCCAACCTGACCT
TGGAAGATGCTGCTGAGTGTCTCAAGCAGCACTGACAGCAGCTGGGCCTGCCCCAGGGCAAC
GTGGGGGCGGAGACTCAGCTGGACAGCCCCCTGCCTGTCACTCTGGAGCTGGGCTGCTGCTGC
CTCAGGACCCCCCTCTCCGACCCCGGACAGAGCTGAGCTGGCCAGGGCCAGGAGGGCGGGAGG
GAGGGAATGGGGGTGGGCTGTGCGCAGCATCAGCGCCTGGGCAGGTCCGCAGAGCTGCGGGA
TGTGATTAAAGTCCCTGATGTTTCTC

FIGURE 208

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399

><subunit 1 of 1, 157 aa, 1 stop

><MW: 17681, pI: 7.65, NX(S/T): 1

MALLLCLVCLTAALAHGCLHCHSNFSSKKFSFYRHHVNFKSWWVGDI PVSGALLTDWSDDTMK
ELHLAIPAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAIIER
HLAPGSWGGGQLSREGPSLAPEGSMPSPRGDLP

Signal peptide:

amino acids 1-15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250

FIGURE 209

AGCAGGAGCAGGAGAGGGACA**ATG**GAAGCTGCCCCGTCCAGGTTTCATGTTCCCTCTTATTTCT
 CCTCACGTGTGAGCTGGCTGCAGAAGTTGCTGCAGAAGTTGAGAAATCCTCAGATGGTCCTG
 GTGCTGCCCAGGAACCCACGTGGCTCACAGATGTCCAGCTGCCATGGAATTCATTGCTGCC
 ACTGAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCA
 TAGCATGGTGC AAAAATCCCAGGCGTGCAATTTGGGATCAGCACTGATTCTGAGGTTCTGA
 CACACTACAACATCACTGGGAACACCATCTGCCTCTTTCGCCTGGTAGACAATGAACAACCTG
 AATTTAGAGGACGAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGAT
 CAACAGCCTCCACATGGTGACAGAGTACAACCCTGTGACTGTGATTGGGTTATTCAACAGCG
 TAATTCAGATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGAGAACATG
 CACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAG
 TGGTATGAAAGAAAATGGGAAGGTGATATCATTTTTCAAATAAAGGAGTCTCAACTGCCAG
 CTTTGGCAATTTACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGCAGAAGTTTCC
 GTAGAGCATGTGCAAACTTTTGTGATGGATTCCCTAAGTGGA AAAATTGTTGAAAGAAAATCG
 TGAATCAGAAGGAAAGACTCCAAAGGTGGAAC**CTGA**CTTCTCCTTGGA ACTACATATGGCC
 AAGTATCTACTTTATGCAAAGTAAAAAGGCACA ACTCAAATCTCAGAGACACTAAACAACAG
 GATCACTAGGCCTGCCAACCACACACACACGACGTGCACACACGCACGCACGCGTGCACAC
 ACACACGCGCACACACACACACACACAGAGCTTCATTTCTGTCTTAAATCTCGTTTTCTC
 TTCTTCCTTCTTTTAAATTTTCATATCCTCACTCCCTATCCAATTTCTTCTTATCGTGCATT
 CATACTCTGTAAGCCCATCTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATG
 CCTCTATGAAAGAGAGGCATTCCCTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGT
 TTGTATACTGCACATGACTTACACACAACATAGTTCCCTGCTCTTTTAAAGGTTACCTAAGGGT
 TGAAACTCTACCTTCTTTTATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGG
 ATGGTTTTAAACACCTTTGTGAAATTGTCTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTC
 CCTGAACTCAGCAGAAATAGACCATGTGAAAACCTCATGCTTGGT TAGCATCTCCA ACTCCC
 TATGTAAATCAACAACCTGCATAATAATAAAAAGGCAATCATGTTATA

FIGURE 210

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30480, pI: 4.60, NX(S/T): 1

MEAAPSRFMFLLFLLTCELAAEVAAEVEKSSDGPAAQEPTWLTDPAAAMEFIAATEVAVIG
FFQDLEIPAVPILHSMVQKFPGVVSFGISTDSEVLTHYNITGNTICLFRRLVDNEQLNLEDEDI
ESIDATKLSRFIEINSLHMTVEYNPVTVIGLFNSVIQIHLLIMNKASPEYEENMHRYQKAA
KLFQKGILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDDEWDTLPTAEVSVEHVQNF
CDGFLSGKLLKENRESEGKTPKVEL

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 143-162

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 211

1

FIGURE 212

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76510

><subunit 1 of 1, 269 aa, 1 stop

><MW: 29082, pI: 9.02, NX(S/T): 3

MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKS
TSTTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDDKSDASINIEN
MQFIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVGIVTAVVLGLTLLISMI
LAVLYRRKNSKRDTGCTSESLSVPKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSG
GHHSKINKSESVYADIRKN

Signal peptide:

amino acids 1-37

Transmembrane domain:

amino acids 161-183

protein = 6865001

FIGURE 213

GCCGGCTGTGCAGAGACGCC**ATG**TACCGGCTCCTGTCAGCAGTGACTGCCCCGGGCTGCCGCC
 CCCGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGCTGCCGCC
 TCTCGGCCACGGCTGGGTGCGGGGCCCTCGGGCTGGGGCTGGGGCTGGCGCTCGGGGTGAAGC
 TGGCAGGTGGGCTGAGGGGCGCGGCCCCGGCGCAGTCCCCCGCGGCCCCCGACCCTGAGGCG
 TCGCCTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGC
 GCCGCCCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGA
 TCAAGGATGAGGTGGGCGCACCGGGCATAGTGGTTGGAGTTTCTGTAGATGGAAAAGAAGTC
 TGGTCAGAAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGT
 TATGCGAATTGCTAGCATCAGCAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAG
 CAGGGAAACTGGATCTTGATATTCCAGTACAACATTATGTTCCCGAATTCCAGAAAAAGAA
 TATGAAGGTGAAAAGGTTTCTGTCAACAAGATTACTGATTTCCCATTTAAGTGGAATTCTG
 TCATTATGAAAAGGACATAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGA
 TGAAAGAGAATGTTGCATTTGAGCAAGAAAAAGAAGGCAAAAGTAATGAAAAGAATGATTTT
 ACTAAATTTAAAACAGAGCAGGAGAATGAAGCCAAATGCCGGAATTCAAAACCTGGCAAGAA
 AAAGAATGATTTTGAACAAGGCGAATTATATTTGAGAGAAAAGTTTGAAAATTCAATTGAAT
 CCCTAAGATTATTTAAAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTTTGTATTCA
 ACTTTTGGCTATAACCCTACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGA
 CTATATGCAGAAAATATTCCATGACTTGGATATGCTGACGACTGTGCAGGAAGAAAACGAGC
 CAGTGATTTACAATAGAGCAAGG**TAA**ATGAATACCTTCTGCTGTGTCTAGCTATATCGCATC
 TTAACACTATTTTATTAATTAAAAGTCAAATTTTCTTTGTTTCCATTCCAAAATCAACCTGC
 CACATTTTGGGAGCTTTTCTACATGTCTGTTTTCTCATCTGTAAAGTGAAGGAAGTAAACA
 TGTTTATAAAGTAAAAAAA

FIGURE 214

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76522
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41221, pI: 8.54, NX(S/T): 0
MYRLLSAVTARAAAPGGLASSCGRRGVHQAGLPPLGHGWVGGLGLGLGLALGVKLAGGLRG
AAPAQSPAAPDPEASPLAEPPEQEQSLAPWSPQTPAPPSCRCFARAIESSRDLLHRIKDEVGA
PGIVVGVSVDGKEVWSEGLGYADVENRVPCPKPETVMRIASISKSLTMVALAKLWEAGKLDLD
IPVQHYVPEFPEKEYEGEKVSVTTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAF
EQEKEGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKN
DPLFFKPGSQFLYSTFGYTLLAAIVERASGCKYLDYMQKIFHDLDMLT TVQEENEPIYNRAR
```

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 39-60

PDB ID: 6EE6

FIGURE 215

GTGACACTATAGAAGAGCTATGACGTGCGATGCACGCGTACGTAAGCTCGGAATTCGGGCTCG
AGGCTGGTGGGAAGAAGCCGAGGATGGCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGC
TCTTGCTGATGGCGGTAGCAGCGCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACT
GGTGCGCGAGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCTGTGGCACGGTGGGGCTGCT
GCTGGAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCT
GGAACCAGCAGGATGGTACCTTGTCCCTGTACAGCGGCAGCTCAGCGAGGAGGAGCGGGGC
CGACTCCGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGG
GGCCCTGGATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGG
TGGAGTCGCACCTGTTCGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGC
GTGTCGGTGGTGACGCACCCCGGGGGCTGCCGGGGGCCATGAGGTGGAGGACGTGGACCTGGA
GCTGTTCAACACCTCGGTGCAGCTGCAGCCGCCACCACAGCCCCAGGCCCTGAGACGGCGG
CCTTCATTGAGCGCTGGAGATGGAACAGGCCCAGAAGGCCAAGAACCCCCAGGAGCAGAAG
TCCTTCTTCGCCAAATACTGGATGTACATCATTC³CGTCGTCCTGTTCTCATGATGTCAGG
AGCGCCAGACACCGGGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCC
TTTGCTGTGTGCCACCCTCCCTGTAAGTCTATTTAAAAACATCGACGATACATTGAAATGTG
TGAACGTTTTTGAAAAGCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGG
TCCTGATGTACAAGCTTGATTGAAATTCCTGCTCACTTGATACGTTATTTCAGAAACCCAAG
GAATGGCTGTCCCCATCCTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTAT
TAAACTGTCCCCCAGATCGACACGCAAAAAAAAAA

FIGURE 216

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76529

><subunit 1 of 1, 269 aa, 1 stop

><MW: 28004, pI: 5.80, NX(S/T): 1

MAAASAGATRLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEI
DDSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGLRDVAALNGLYRVRIPRRPGALDGLEA
GGYVSSFVPACSLVESHLSDDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELFNTSVQ
LQPPTTAPGPETAFAFIERLEMEQAQKAKNPQEQKSFFAKYWMYIIPVVLFLMMSGAPDTGGQ
GGGGGGGGGGGSGLCVPPSL

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 226-243

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 217

GGAGCGCTGCTGGAACCCGAGCCGGAGCCGGAGCCACAGCGGGGAGGGTGGCCTGGCGGCCT
 GGAGCCGGACGTGTCCGGGGCGTCCCCGCAGACCGGGGCAGCAGGTGCTCCGGGGGCCACC
ATGCTGGTGACTGCCTACCTTGCTTTTGTAGGCCTCCTGGCCTCCTGCCTGGGGCTGGAAC
 GTCAAGATGCCGGGCTAAACCCCTGGAAGGGCCTGCAGCAATCCCTCCTTCCTTCGGTTTC
 AACTGGACTTCTATCAGGTCTACTTCCTGGCCCTGGCAGCTGATTGGCTTCAGGCCCCCTAC
 CTCTATAAACTCTACCAGCATTACTACTTCCTGGAAGGTCAAATTGCCATCCTCTATGTCTG
 TGGCCTTGCCCTCTACAGTCCTCTTTGGCCTAGTGGCCTCCTCCCTTGTGGATTGGCTGGGTC
 GCAAGAATTCTTGTGTCTCTTCTCCCTGACTTACTCACTATGCTGCTTAACCAAACCTCTCT
 CAAGACTACTTTGTGTGCTAGTGGGGCGAGCACTTGGTGGGCTGTCCACAGCCCTGCTCTT
 CTCAGCCTTCGAGGCCTGGTATATCCATGAGCACGTGGAACGGCATGACTTCCCTGCTGAGT
 GGATCCCAGCTACCTTTGCTCGAGCTGCCTTCTGGAACCATGTGCTGGCTGTAGTGGCAGGT
 GTGGCAGCTGAGGCTGTAGCCAGCTGGATAGGGCTGGGGCCTGTAGCGCCCTTTGTGGCTGC
 CATCCCTCTCCTGGCTCTGGCAGGGGCCTTGGCCCTTCGAAACTGGGGGGAGAACATATGACC
 GGCAGCGTGCCCTTCTCAAGGACCTGTGCTGGAGGCCTGCGCTGCCTCCTGTCGGACCGCCGC
 GTGCTGCTGCTGGGCACCATAACAAGCTCTATTTGAGAGTGTGTCATCTTCATCTTTGTCTTCCT
 CTGGACACCTGTGCTGGACCCACACGGGGCCCCCTCTGGGCATTATCTTCTCCAGCTTCATGG
 CAGCCAGCCTGCTTGGCTCTTCCCTGTACCGTATCGCCACCTCCAAGAGGTACCACCTTCAG
 CCCATGCACCTGCTGTCCCTTGCTGTGCTCATCGTCGTCTTCTCTCTTTCATGTTGACTTT
 CTCTACCAGCCCAGGCCAGGAGAGTCCGGTGGAGTCCCTCATAGCCTTTCTACTTATTGAGT
 TGGCTTGTGGATTATACTTTCCCAGCATGAGCTTCCTACGGAGAAAGGTGATCCCTGAGACA
 GAGCAGGCTGGTGTACTCAACTGGTTCCGGGTACCTCTGCACTCACTGGCTTGCCTAGGGCT
 CCTTGTCCTCCATGACAGTGATCGAAAAACAGGCACTCGGAATATGTTTCAGCATTTGCTCTG
 CTGTCATGGTGATGGCTCTGCTGGCAGTGGTGGGACTCTTCACCGTGGTAAGGCATGATGCT
 GAGCTGCGGGTACCTTCACCTACTGAGGAGCCCTATGCCCTGAGCTG**TAA**CCCCACTCCAG
 GACAAGATAGCTGGGACAGACTCTTGAATTCCAGCTATCCGGGATTGTACAGATCTCTCTGT
 GACTGACTTTGTGACTGTCTGTGGTTTCTCCTGCCATTGCTTTGTGTTGGGAGGACATGA
 TGGGGGTGATGGACTGGAAAGAAGGTGCCAAAAGTTCCTCTGTGTTACTCCCATTTAGAAA
 ATAAACACTTTTAAATGATCAAAAAAAAAA

FIGURE 218

MLVTAYLAFVGLLASCLGLELSRCRAKPPGRACSNPSFLRFQLDFYQVYFLALAADWLQAPY
LYKLYQHYYFLEGQIAILYVCGLASTVLFGLVASSLVDWLGRKNSCVLFSLTYSLCCLTKLS
QDYFVLLVGRALGGLSTALLFSAFEAWYIHEHVERHDFPAEWIPATFARAAFWNHVLAVVAG
VAAEAVASWIGLGPVAPFVAAIPLLALAGALALRNWGENYDRQRAFSRTCAGGLRCLLSDRR
VLLLGTIQALFESVIFIFVFLWTPVLDPHGAPLGII FSSFMAASLLGSSLYRIATSKRYHLQ
PMHLLSLAVLIVVFSLEMLTFSTSPGQESPVESFIAFLLIELACGLYFPSMSFLRRKVIPET
EQAGVLNWFRVPLHSLACLGLLVLHDSDRKTGTRNMF SICSAVMVMALLAVVGLFTTVVRHDA
ELRVPSPTTEEPYAPEL

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 41-55, 75-94, 127-143, 191-213, 249-270, 278-299,
314-330, 343-359, 379-394, 410-430

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 219

GCGACGCGCGGGCGGGGCGGCGAGAGGAAACGCGGCGCCGGGCCGGGCCCGGCCCTGGAG**ATG**
 GTCCCCGGCGCCGCGGGCTGGTGTGTCTCGTGCTCTGGCTCCCCGCGTGCGTCGCGGCCCA
 CGGCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTCGATACA
 TCTTCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATGAGCAGATT
 CACCTTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACCTCAGCAACGGTTTCTTCATCCA
 GGACCAGATTGCTCTGGTGGAGAGGGGGGGCTGCTCCTTCCTCTCCAAGACTCGGGTGGTCC
 AGGAGCACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTAC
 GTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCCTGCTCGG
 CCGAGACGGCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTT
 CCATCCCAGTCAATGTCACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTC
 TGG**TAG**AAGAGTTTGTCCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCC
 AGGAATTTTGCTACTTGGAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGA
 AAAGGGTTTGGGCGTTGCTAGGCTGAAAGGGAAGCCACACCACTGGCCTTCCCTTCCCCAGG
 GCCCCAAGGGTGTCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTA
 GAACCCGAAACAAAAGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACT
 CACCTGGCTCCAGCCTCCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGG
 AGTGGTTTAAAGAGCTGGTGTGTTGSGGACTCAATAAACCTCACTGACTTTTGTAGCAATAAA
 GCTTCTCATCAGGGTTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 220

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532
```

><subunit 1 of 1, 188 aa, 1 stop

><MW: 21042, pI: 5.36, NX(S/T): 2

MVPGAAGWCCLVLWLPACVAAHGFRIDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQ
IHLVPAEPPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIIISDNAVDNDSF
YVEMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW

Signal peptide:

amino acids 1-20

FIGURE 221

TCTGCCTCCACTGCTCTGTGCTGGGATCATGGAACTTGCACTGCTGTGTGGGCTGGTGGTGA
TGGCTGGTGTGATTCCAATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTG
ACTGGGAAAATGCCCATCCTCTCCTACTGGCCCTACGGCTGTCACTGCGGACTAGGTGGCAG
AGGCCAACCCAAAGATGCCACGGACTGGTGCTGCCAGACCCATGACTGCTGCTATGACCACC
TGAAGACCCAGGGGTGCGGCATCTACAAGGACAACAACAAAAGCAGCATACATTGTATGGAT
TTATCTCAACGCTATTGTTTAATGGCTGTGTTTAATGTGATCTATCTGGAAAATGAGGACTC
CGAATAAAAAGCTATTACTAWTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 222

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76538
><subunit 1 of 1, 116 aa, 1 stop
><MW: 12910, pI: 6.41, NX(S/T): 1
MELALLCGLVVMAGVPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDW
CCQTHDCCYDHLKTQGCGLYKDNKSSIHCMDSLQRYCLMAVFNVIYLENEDSE
```

Important features of the protein:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 1-24

N-glycosylation site.

amino acids 86-89

N-myristoylation sites.

amino acids 20-25, 45-50

Phospholipase A2 histidine active site.

amino acids 63-70

2025-03-01 10:00:00

FIGURE 223

CTCGCTTCTTCCTTCTGGATGGGGGCCCAGGGGGCCCAGGAGAGTATAAAGGCGATGTGGAG
GGTGGCCCGGCACAACCAGACGCCCAGTCACAGGCGAGAGCCCTGGG**ATG**CACCGGCCAGAGG
CCATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTAT
GGCCCTGGAGGAGGCAAGTATTTACGACCACTGAAGACTACGACCATGAAATCACAGGGCT
GCGGGTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCAGGTGAAACTTGGAGACTCCTGGG
ACGTGAAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATAC
ATCACAAAAGTCTTTGTCGCCTTCCAAGCTTTCCTCCGGGGTATGGTCATGTACACCAGCAA
GGACCGCTATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCCAGCCAAG
AGGGGCAGGTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGC
TTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCACCAGTTAATCTCACATACTC
AGCAAACTCACCCGTGGGTCGCT**TAG**GGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGT
GTGGTGGTGGCTGATGGTACTGGAGTAACTGAGTCGGGACGCTGAATCTGAATCCACCAATA
AATAAAGCTTCTGCAGAAAA

FIGURE 224

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
```

><subunit 1 of 1, 178 aa, 1 stop

><MW: 19600, pI: 5.89, NX(S/T): 1

MHRPEAMLLLLLTALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKSVQVK
LGDSWDVKLGALGGNTQEVTLQPGEYITKVFFAFQAFLRGMVMYTSKDRYFYFGKLDGQISS
AYPSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR

Signal peptide:

amino acids 1-22

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCT
GAACTGGGTGCTCATCACGGGAAGCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAG
CCCCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTT
TTTTTTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAAGCTGTAAAGATGCAAAAACGTAATAT
CCATGAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATT
TATTTGTTCTTGGAGTGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCT
CCCAAGGGGTCCAATTTTTCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTG
ACAGGGGCTGTCATGCAACTGGCCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAA
CAATACAAAGGATGGGTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTT
ATAGCCCCCACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTG
TAGGTGTGAAGGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCCTCAAGTA
TATCTGCTGGTTGCTTAGGTTTTGTCCCTTCGCTATAACAGCCTTCAAAAACCTAAGTATAAT
CAATTTAAAGGGCTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATAT
TGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAA
TCTCCTATTTTTCTTAACAATACCTTCAGACCTGTGACAAATTTACGGAAGCTTGGATCTGTCC
TATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTTCGGGGCTTGCGGAAGCTGCTGAGTTT
ACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACC
TGGAAGTTTTTGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGC
ATGATCAGACTCAAAGAAGCTCACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCT
TTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTCATAG
GACAGACCATGTCCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATC
GAAGCTTTTCAGTGGAACCCAGTGTTTTCCAGTGTTGCCGAATCTGCAGCGCCTCAACCTGGA
TTCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGATATCCCTCAATGACA
TCAGTCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGCTGGCTG
AAAAGTTTTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGG
AGTAAATGTGATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGT
TTGATCTGGCCAGGGCTCTCCCAAAGCCGACGTTTAAAGCCCAAGCTCCCCAGGCCGAAGCAT
GAGAGCAAACCCCTTTGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGC
TGACGCCGAGCACATCTTTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCG
TGCTCGTCATCTGCTGTTTATCTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAG
CTGCAGCAGCGCTCCCTCATGCGAAGGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAAT
GACTCCCAGCACCAGGAATTTTATGTAGATTATAAACCACCAACACGGAGACCAGCGAGA
TGCTGCTGAATGGGACGGGACCCCTGCACCTATAACAAATCGGGCTCCAGGGAGTGTGAGGTA
TGAACCATTGTGTATAAAAAGAGCTCTTAAAGAGCTGGGAAATAAGTGGTGTTTATTGAAGCT
TGGTGACTATCAAGGGAACGCGATGCCCCCCCTCCCCTTCCCTCTCCCTCTCACTTTGGTGG
CAAGATCCCTTCCTTGTCGGTTTTAGTGCAATTCATAATAACTGGTCATTTTCCCTCTCATACATA
ATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTTGAAGCTCCGGTTTTAATATAA
TACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTCGCATTTGTTTTAAGATAAAACT
TCTTTCATAGGTAAAAA

FIGURE 226

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77301

><subunit 1 of 1, 513 aa, 1 stop

><MW: 58266, pI: 9.84, NX(S/T): 4

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCCEGKMVYCESQKLQEIPSSISAG
CLGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYF
LNNTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELL
DLGYNRIRSLARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTM
SWTWSSSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIQEILDSWISLNDISLA
GNIWECSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLA
RALPKPTFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVI
LLVIYVSWKRYPASMKQLQQRSLMRRHRKKKRQSLKQMTPTSTQEFYVDYKPTNTTETSEMLLN
GTGPCTYNKSGSRECEV

Important features of the protein:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 227

AGTTCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTT
TAAATATGTCAAGATCCAGACTTTTCAGTGTCACCTCAGCGATCTCAACGATAGGGATCTTG
TGTTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCGAAGAAGATGAAATGTGTGT
AAATTATAATGACCAACACCCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTGG
TGGCAGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGA
ATTGATTCTCACAGGCGCACCATGGCAGTTTTTGTCTGTTGGAGACTTGGACTCTATTTATGG
GACAGAAGCAGCTGTGAGTCCAACCTGTTGGAATTCACCTTCAAACCTCAAACCCCTGACCTAT
ATCCTGTTCCCTGCTCCATGTTTTGGCCCTTTAGGCTCCCCACCTCCATATGAAGAAATTGTA
AAAACAACCTTGATTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCCA
AAACATCAAATTTAGGAATAGTTATTTTCAGTTGTTGGAAATGTCCAGAGATCTATTCATATA
GTCTGAGGAAGGACAATTCGACAAAAGAATGGATGTTGGAAAAAATTTTGGTCATGGAGATG
TTTAAATAGTAAAGTAGCAGGCTTTTGATGTGTCACTGCTGTATCATACTTTTATGCTACAC
AACCAAATTAATGCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTT
TCCATCACATTTAGGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTC
CTAGCATGGGGTCCATAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACAT
GTCCAGAACCAGAACCAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAA
GAAAAGTTTGGAGTTGAAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCT
CAGCAAAAACAAGAGGTTTTATGCCCCAACCTGAAGAGGAAGAAATTGTAGATAGAAGGTGA
AGGAGATTGCTGAAGATATAGAGCACATATAATGCCAACACGGGGAGAAAAGAAAATTTCCC
CTTTTACAGTAATGAATGTGGCCTCCATAGTCCATAGTGTCTCTGGAGCCTCAGGGCTTG
GCATTTATTGCAGCATCATGCTAAGAACCTTCGGCATAGGTATCTGTTCCCATGAGGACTGC
AGAAGTAGCAATGAGACATCTTCAAGTGGCATTGTTGGCAGTGGCCATCAGCAGGGGGACAGA
CAAAAACATCCATCACAGATGACATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAAT
AAACATCAATAGATATCTAAAAA

FIGURE 228

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

><subunit 1 of 1, 146 aa, 1 stop

><MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSAISTIGILCLPLFQLVLSDLPCCEEDEMCVNYNDQHPNGWYIWILLLLVLVA
ALLCGAVVLCLQCWLRRPRIDSHRRTMVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYP
VPAPCFGPLGSPPPYEEIVKTT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 52-70

Transmembrane domain:

FIGURE 229

GAGCGGAGTAAAATCTCCACAAGCTGGGAACAAACCTCGTCCCAACTCCCACCCACCGGCGT
 TTCTCCAGCTCGATCTGGAGGCTGCTTCGCCAGTGTGGGACGCAGCTGACGCCCCTTATTA
 GCTCTCGCTGCGTCGCCCCGGCTCAGAAGCTCCGTGGCGGCGGCGACCGTGACGAGAAGCCC
 ACGGCCAGCTCAGTTCTCTTCTACTTTGGGAGAGAGAGAAAGTCAGATGCCCCCTTTTAACT
 CCCTCTTCAAACTCATCTCCTGGGTGACTGAGTTAATAGAGTGGATAACAACCTTGCTGAAG
 ATGAAGAATATACAATATTGAGGATATTTTTTTCTTTTTTTTTTCAAGTCTTGATTTGTGGC
 TTACCTCAAGTTACCATTTTTTCAGTCAAGTCTGTTTGTGTTGCTTCTTCAGAAATGTTTTTTA
 CAATCTCAAGAAAAAATATGTCCAGAAATTGAGTTTACTGTTGCTTGTATTTGGACTCATT
 TGGGGATTGATGTTACTGCACTATACTTTTCAACAACCAAGACATCAAAGCAGTGTCAAGTT
 ACGTGAGCAAATACTAGACTTAAGCAAAGATATGTTAAAGCTCTAGCAGAGGAAAATAAGA
 ACACAGTGGATGTCGAGAACGGTGCTTCTATGGCAGGATATGCGGATCTGAAAAGAACAATT
 GCTGTCCTTCTGGATGACATTTTGAACGATTGGTGAAGCTGGAGAACAAAGTTGACTATAT
 TGTTGTGAATGGCTCAGCAGCCAACACCACCAATGGTACTAGTGGGAATTTGGTGCCAGTAA
 CCACAAATAAAAGAACGAATGTCTCGGGCAGTATCAGATAGCAGTTGAAAATCACCTTGTGC
 TGCTCCATCCACTGTGGATTATATCCTATGGCAGAAAAGCTTTATAATTGCTGGCTTAGGAC
 AGAGCAATACTTTACAATAAAAGCTCTACACATTTTCAAGGAGTATGCTGGATTCATGGAAC
 TCTAATTCTGTACATAAAAAATTTTAAAGTTATTTGTTTGTCTTCAGGCAAGTCTGTTCAATG
 CTGTACTATGTCCTTAAAGAGAATTTGGTAACTTGGTTGATGTGGTAAGCAGATAGGTGAGT
 TTTGTATAAATCTTTTGTGTTTGAAGATCAAGCTGAAATGAAAACACTGAAAAACATGGATTC
 ATTTCTATAACACATTTATTTAAGTATATAACACGTTTTTTGGACAAGTGAAGAATGTTTAA
 TCATTCTGTCAATTTGTTCTCAATAGATGTAAGTGTAGACTACGGCTATTTGAAAAAATGTG
 CTTATTGTACTATATTTTGTATTCCAATTATGAGCAGAGAAAGGAAATATAATGTTGAAAA
 TAATGTTTTGAAATCATGACCCAAAGAATGTATTGATTTGCACTATCCTTCAGAATAACTGA
 AGGTTAATTATTGTATATTTTTTAAAAATTACACTTATAAGAGTATAATCTTGAAATGGGTAG
 CAGCCACTGTCCATTACCTATCGTAAACATTTGGGGCAATTTAATAACAGCATTAAATAGTT
 GTAAACTCTAATCTTATACTTATTGAAGAATAAAAGATATTTTTATGATGAGAGTAACAATA
 AAGTATTCATGATTTTTTCACATACATGAATGTTTCAATTTAAAGTTTAAATCCTTTGAGTGTCT
 ATGCTATCAGGAAAGCACATTATTTCCATATTTGGGTAAATTTTGCTTTTATTATATTGGTC
 TAGGAGGAAGGGACTTTGGAGAATGGAACCTTGAGGACTTTAGCCAGGTGTATATAATAAA
 GGTACTTTTGTGCTGCATTAAATTGCTTGGAAGTGTTAACATTATATTATATAAGAGTATC
 CTTTATGAAATTTTGAATTTGTATAACAGATGCATTAGATATTCATTTTATATAATGGCCAC
 TTAAAAATAAGAACATTTAAAAATATAAACTATGAAGATTGACTATCTTTTCAGGAAAAAAGCT
 GTATATAGCACAGGGAACCCTAATCTTGGGTAAATCTAGTATAAAACAAATTATACTTTTAT
 TTAAATTTCCCTTGTAGCAAATCTAATTGCCACATGGTGCCCTATATTTTCATAGTATTTATT
 CTCTATAGTAACTGCTTAAAGTGCAGCTAGCTTCTAGATTTAGACTATATAGAATTTAGATAT
 TGTATTGTTTCGTCATTATAATATGCTACCACATGTAGCAATAATTACAATATTTTATTAAAA
 TAAATATGTGAAATATTGTTTCATGAAAGACAGATTTCCAAATCTCTCTTCTCTCTCTGTA
 CTGTCTACCTTTATGTGAAGAAATTAATTATATGCCATTGCCAGGT

FIGURE 230

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648
><subunit 1 of 1, 140 aa, 1 stop
><MW: 15668, pI: 10.14, NX(S/T): 5
MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE
ENKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNL
VPVTTNKRTNVSGSIR
```

Important features of the protein:

Signal peptide:

amino acids 1-26

FIGURE 231

CGCGGCCGGGCCGCCGGGGTGAGCGTGCCGAGGCGGCTGTGGCGCAGGCTTCCAGCCCCAC
CATGCCGTGGCCCCTGCTGCTGCTGCTGGCCGTGAGTGGGGCCCAGACAACCCGGCCATGCT
TCCCCGGGTGCCAATGCGAGGTGGAGACCTTCGGCCTTTTCGACAGCTTCAGCCTGACTCGG
GTGGATTGTAGCGGCCCTGGGCCCCCACATCATGCCGGTGCCCATCCCTCTGGACACAGCCCA
CTTGGACCTGTCCTCCAACCGGCTGGAGATGGTGAATGAGTCGGTGTGGCGGGGCCGGGCT
ACACGACGTTGGCTGGCCTGGATCTCAGCCACAACCTGCTCACCAGCATCTCACCCACTGCC
TTCTCCCGCCTTCGCTACCTGGAGTCGCTTGACCTCAGCCACAATGGCCTGACAGCCCTGCC
AGCCGAGAGCTTCACCAGCTCACCCCTGAGCGACGTGAACCTTAGCCACAACCAGCTCCGGG
AGGTCTCAGTGTCTGCCTTCACGACGCACAGTCAGGGCCGGGCACTACACGTGGACCTCTCC
CACAACCTCATTCACCGCCTCGTGCCCCACCCACAGAGGGCCGGCCTGCCTGCGCCCACCAT
TCAGAGCCTGAACCTGGCCTGGAACCGGCTCCATGCCGTGCCAACCTCCGAGACTTGCCCC
TGCGCTACCTGAGCCTGGATGGGAACCTCTAGCTGTCATTGGTCCGGGTGCCTTCGCGGGG
CTGGGAGGCCTTACACACCTGTCTCTGGCCAGCCTGCAGAGGCTCCCTGAGCTGGCGCCCAG
TGGCTTCCGTGAGCTACCGGGCCTGCAGGTCTTGACCTGTCGGGCAACCCCAAGCTTAAC
GGGCAGGAGCTGAGGTGTTTTTTCAGGCCTGAGCTCCCTGCAGGAGCTGGACCTTTCGGGCACC
AACCTGGTGCCCCCTGCCTGAGGCGCTGCTCCTCCACCTCCCGGCACTGCAGAGCGTCAGCGT
GGGCCAGGATGTGCGGTGCCGGCGCCTGGTGCGGGAGGGCACCTACCCCGGAGGCCTGGCT
CCAGCCCCAAGGTGCCCTGCACTGCGTAGACACCCGGGAATCTGCTGCCAGGGGGCCCCACC
ATCTT**GTGA**CAAATGGTGTGGCCAGGGCCACATAACAGACTGCTGTCTGGGCTGCCTCAG
GTCCCGAGTAACTTATGTTCAATGTGCCAACACCAGTGGGGAGCCCGCAGGCCTATGTGGCA
GCGTCACCACAGGAGTTGTGGGCCTAGGAGAGGCTTTGGACCTGGGAGCCACACCTAGGAGC
AAAGTCTCACCCCTTTGTCTACGTTGCTTCCCCAAACCATGAGCAGAGGGACTTCGATGCCA
AACCAGACTCGGGTCCCCTCCTGCTTCCCTTCCCCACTTATCCCCAAGTGCCTTCCCTCAT
GCCTGGGCGGGCCTGACCCGCAATGGGCAGAGGGTGGGTGGGACCCCTGCTGCAGGGCAGA
GTTTCAGGTCCACTGGGCTGAGTGTCCCTTGGGCCCATGGCCAGTCACTCAGGGGCGAGTT
TCTTTTCTAACATAGCCCTTTCTTTGCCATGAGGCCATGAGGCCCGCTTCATCCTTTTCTAT
TTCCCTAGAACCTTAATGGTAGAAGGAATTGCAAAGAATCAAGTCCACCCTTCTCATGTGAC
AGATGGGGAAACTGAGGCCTTGAGAAGGAAAAAGGCTAATCTAAGTTCCTGCGGGCAGTGGC
ATGACTGGAGCACAGCCTCCTGCCTCCCAGCCCCGACCCAATGCACTTTCTTGTCTCCTCTA
ATAAGCCCCACCCTCCCCGCTGGGCTCCCCTTGCTGCCCTTGCTTCCCCATTAGCACA
GGAGTAGCAGCAGCAGGACAGGCAAGAGCCTCACAAGTGGGACTCTGGGCCTCTGACCAGCT
GTGCGGCATGGGCTAAGTCACTCTGCCCTTCGGAGCCTCTGGAAGCTTAGGGCACATTGGTT
CCAGCCTAGCCAGTTTCTCACCCCTGGGTTGGGGTCCCCCAGCATCCAGACTGGAAACCTACC
CATTTTCCCCTGAGCATCCTCTAGATGCTGCCCCAAGGAGTTGCTGCAGTTCTGGAGCCTCA
TCTGGCTGGGATCTCCAAGGGGCCTCCTGGATTTCAGTCCCCACTGGCCCTGAGCACGACAGC
CCTTCTTACCCTCCCAGGAATGCCGTGAAAGGAGACAAGGTCTGCCCCACCCATGTCTATGC
TCTACCCCCAGGGCAGCATCTCAGCTTCCGAACCTGGGCTGTTTCCTTAGTCTTCATTTTA
TAAAAGTTGTTGCCTTTTAAACGGAGTGTCACTTTCAACCGGCCCTCCCCTACCCCTGCTGGC
CGGGGATGGAGACATGTCATTTGTAAAAGCAGAAAAAGGTTGCATTTGTTCACTTTTGTAAT
ATTGTCCTGGGCCTGTGTTGGGGTGTGGGGGAAGCTGGGCATCAGTGGCCACATGGGCATC
AGGGGCTGGCCCCACAGAGACCCACAGGGCAGTGAGCTCTGTCTTCCCCACCTGCCTAGC
CCATCATCTATCTAACCGGTCCTTGATTTAATAAACACTATAAAAAGGTTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77652

><subunit 1 of 1, 353 aa, 1 stop

><MW: 37847, pI: 6.80, NX(S/T): 2

MPWPLLLLLAVSGAQTTTRPCFPGCQCEVETFGLFDSFSLTRVDCSGLGPHIMPVPIPLDTAH
LDLSSNRLEMVNESVLGPGYTTLAGLDLSHNLLTSISPTAFSRLRYLESLDLSHNGLTALP
AESFTSSPLSDVNLSHNQLREVSVSFAFTTHSQGRALHVDLSHNLIHRLVPHPTRAGLPAPTI
QSLNLAWNRLHAVPNLRDLPLRYLSLDGNPLAVIGPGAFAFAGLGGLTHLSLASLQRLPELAPS
GFRELPGLOVLDLSGNPKLNWAGAEVFSGLSSLQELDLSGTNLVPLPEALLLHLPALQSVSV
GQDVRCRRLVREGTYPRRPGSSPKVPLHCVDTRESAARGPTIL

Signal peptide:

amino acids 1-16

Transmembrane domains:

amino acids 215-232, 287-304

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 233

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTTCTCCCCAGTTCCCCCTGTGGGTCTGAGC
GGACCAGAAGGGTGAGCTACGTTGGCTTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCC
AAACAAGTTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCT
GTTCCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGC
CACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCA
GGGAGGCTTGGCAGTTTTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCC
TCTAGTCTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGG
ACTGAAGACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATG
GATTTTCTGAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTA
AGGAGGACTGAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTT
GCTAAGACTCTATCTGGACAGGGTATTTTAAAACTACCAGACCCCTGACCATTATACTCTCC
GGAAGATCAGCAGCCTCGCCAATTCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCAT
GCCCACATGACATGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCA
CTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTC
TGCAATGGATGGAGGAGACAGAATTAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGT
CAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACT
GTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGAT
TGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTTTTGTAATATCTT
TCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTT
ATTTTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAG
AGCAGGTGATGTATTTTTTATACAGTAAAAAATAAACCTTGTAATTTCTAGAAGAGTGGCT
AGGGGGGTTATTCAATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGA
TATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATT
GCACATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTAT
CTTCCAGCCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATA
CCAAAAAAAAAAAAAAAAAAAA

FIGURE 234

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500
```

```
><subunit 1 of 1, 261 aa, 1 stop
```

><MW: 29667, pI: 8.76, NX(S/T): 0

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTWCWALTAEPGWGQNKGATTCATNSHSDSELRP
EIFSSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNL
QEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTP
DHYTLRKISSLANSFLTIIKKDLRLSHAHMTCHCGEAMKKYSQILSHFEKLEPQAAVVKALG
ELDILLQWMEETE

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 235

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCG
CTGACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGC
CGAGCTAGCAACCTTTCCCCTGGATCTCACAAAACCTCGACTCCAAATGCAAGGAGAAGCAG
CTCTTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACA
GCCCTAGGGATCATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCAT
TTACAGACACGTAGTGTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTG
TGTTTGCCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATG
GCTGGTGTTATTGGCCAGTTTTTAGCCAATCCAACTGACCTAGTGAAGGTTGAGATGCAAAT
GGAAGGAAAAAGGAACTGGAAGGAAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTG
CAAAAATCTTAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAA
AGAGCAGCACTGGTGAATATGGGAGATTTAACCACCTTATGATACAGTGAAACACTACTTGGT
ATTGAATACACCACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGAC
TGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCA
CGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTGAGGCTGT
TCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACCATCTTGGCTGAGAATGACCC
CTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGATGAGTGGAGTCAGTCCA
TTT**TAA**

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77568

><subunit 1 of 1, 323 aa, 1 stop

><MW: 36064, pI: 9.33, NX(S/T): 1

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGA
RESAPYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEH
YPLWKSVIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGI
RGLWAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTP
ADVIKSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLT
YEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 237

CGGACGCGTGGGCGCGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGC
GCCTGAAGTCGGCGTGGGCGTTTGAGGAAGCTGGGATACAGCATTTAATGAAAAATTTATGC
TTAAGAAGTAAAA**ATG**GCAGGCTTCCTAGATAATTTTCGTTGGCCAGAATGTGAATGTATTG
ACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTGCGCAGGTATATTGTTTTTTACAGGC
TGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGAACCATGCCTT
TCACACATGTGGTGTATTTTCCACATTGGCTTTCTTCATGATAAATGCTGTATCCAATGCTC
AGGTGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTT
TTCATTGGTTTCATGTTGATGTTTGGGTCACTTATTGCTTCCATGTGGATTCTTTTTTGGTGC
ATATGTTACCCAAAATACTGATGTTTATCCGGGACTAGCTGTGTTTTTTCAAATGCACTTA
TATTTTTTTAGCACTCTGATCTACAAATTTGGAAGAACCGAAGAGCTATGGACCT**AG**ATCAC
TTCTTAAGTCACATTTTCTTTTGTATATTCTGTTTGTAGATAGGTTTTTTATCTCTCAGT
ACACATTGCCAAATGGAGTAGATTGTACATTAAATGTTTTGTTTCTTTACATTTTTATGTTT
TGAGTTTGAAGTAGTTTTATGAAATTTCTTTATTTTTTCATTGCATAGACTGTTAATATGA
TATAATACAAGACTATATGAATTGGATAATGAGTATCAGTTTTTTATTCTGAGATTTAGAA
CTTGATCTACTCCCTGAGCCAGGGTTACATCATCTTGTCATTTTAGAAGTAACCACTCTTGT
CTCTCTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGG
CCGATTGCTTGAGGTCAAGTGTGTTGAGACCAGCCTGGCCAACATGGCGAAACCCCATCTACT
AAAAATACAAAATTAGCCAGGCATGGTGGTGGGTGCCTGTAATCCCAGCTACCTGGGAGGC
TGAGGCAGGAGAATCGCTTGAACCCGGGGGGCAGAGGTTGCAGTGAGCTGAGTTTGCGCCAC
TGCACTCTAGCCTGGGGGAGAAAGTGAACTCCCTCTCAAAAAAAGACCACTCTCAGTATC
TCTGATTTCTGAAGATGTACAAAAAATATAGCTTCATATATCTGGAATGAGCACTGAGCCA
TAAAAGGTTTTTCAGCAAGTTGTAACCTATTTTGGCCTAAAAATGAGGTTTTTTTGGTAAAGA
AAAAATATTTGTTCTTATGTATTGAAGAAGTGACTTTTATATAATGATTTTTTAAATGCCC
AAAGGACTAGTTTGAAAGCTTCTTTTAAAAAGAATTCCTCTAATATGACTTTATGTGAGAA

FIGURE 238

MAGFLDNFRWPECECIDWSERRNAVASVVAGILFFTGWIMIDAADVYPKPEQLNHAFHTCG
VFSTLAFFMINAVSNAQVRGDSYESGCLGRTGARVWLFIFGMLMFGSLIASMWILFGAYVTQ
NTDVYPGLAVFFQNALIFFSTLIYKFG RTEELWT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 23-42 (type II), 60-80, 97-117, 128-148

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 239

TTGTATGGCAAACCTTCCTCAAAGGAGGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCCA
CTGGCGGCCCCGCAACACTCCGTCTCACCTCTGGGCCCACTGCATCTAGAGGAGGGCCGTCT
GTGAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGT
GGTCAGCTGGGTCAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAG
GGAAGTGGGAGCCTCGAGCCCTCGGGTGGAAGCTGACCCCAAGCCACCTTCACCTGGACAG
GATGAGAGTGTCAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGA
TGTTTATTTCGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCA
GCCTCGCCACCAAGGAGATCCAGGTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCC
AGCCAATACTTTGCGTTTAAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGT
GCTTTGAAGACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATC
GCCCTGGTGAATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGG
AGATGTTATGCACCTAGTGAAATTCCTTAAAGAAATTCCGGGGGGTGCCTGGTGCTGGTGG
CCTCCTACGACGATCCAGGGACCAAAATGAACGATGAAAGCAGGAAACTCTTCTCTGACTTG
GGGAGTTCCTACGCAAAACAACTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGA
CCTCAGGGGTAAAAGCCCCCTTTGAGCAGTTCTTAAAGAACAGCCAGACACAAACAAATACG
AGGGATGGCCAGAGCTGCTGGAGATGGAGGGCTGCATGCCCCGAAGCCATTTTTAGGGTGGC
TGTGGCTCTTCCTCAGCCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCC
GCAGGGGCTGAGGAGGAGGAGCAGGGGGTGCTGCGTGGAAGGTGCTGCAGGTCCTTGCACGC
TGTGTCGCGCCTCTCCTCCTCGGAAACAGAACCTCCACAGCACATCCTACCCGGAAGACC
AGCCTCAGAGGGTCTTCTGGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTTCGTAGGG
ACTGCTGACGGCTGGTCTGAGGAAGGACAACTGCCCAGACTTGAGCCCAATTAAATTTTA
TTTTTGCTGGTTTTTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 240

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59814

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24963, pI: 9.64, NX(S/T): 1

MRVSGVLRLLALIFAIVTTWMFIRSYMSFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCP
ANYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSG
DVMHLVKFLKEIPGGALVLVASYDDPGTKMNDESRKLFSDLGSSYAKQLGFRDSWVFIGAKD
LRGKSPFEQFLKNSPDTNKYEGWPELLEMEGCMPPKPF

Important features:**Signal peptide:**

amino acids 1-15

ATP/GTP-binding site motif A (P-loop).

amino acids 184-191

N-glycosylation site.

amino acids 107-110

protein database

FIGURE 241

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGAT
 CCAGAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACA
 GCAAGAGAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTC
 TCACTCCTCCCTCCCTCTCTCTCTGCTGTCTAGTCCTCTAGTCCTCAAATCCCAGTCCC
 CTGCACCCCTTCCTGGGACACT**ATG**TTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCC
 TGGCTGCAGATGGGGGTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGGCCA
 GCCTCTTACCCTGAGTGTGGAAACAATGCCAGTCGCCATCGATATTCAGACAGACAGTGT
 GACATTTGACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGC
 CTTTGGACCTGCACAACAATGGCCACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTG
 GGTGGACTTCCCCGAAAATATGTAGCTGCCAGCTCCACCTGCACTGGGGTCAGAAAGGATC
 CCCAGGGGGGTGAGAACCAGATCAACAGTGAAGCCACATTTGCAGAGCTCCACATTGTAC
 ATTATGACTCTGATTCCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCT
 GTCCTGGGCATCCTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAG
 TCACTTGCATGAAGTCAGGCATAAAGATCAGAAGACCTCAGTGCCTCCCTTCAACCTAAGAG
 AGCTGCTCCCCAAACAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACCTCCCCCT
 TGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAGATTTCAATGGAACAGCT
 GGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGA
 ACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTTCATCCAAGCAGGA
 TCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGTTGGCTGTCTCTG
 CCTTCTCCTGGCTGTTTATTTTATTGCTAGAAAGATTCGGAAGAAGAGGCTGGAAAACCGAA
 AGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCA**TAA**ATTCCTTCTCAGATACCA
 TGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTG
 GCCAGAAACACTGTAGGAGTAGTAAGCAGATGTCTCCTTCCCCTGGACATCTCTTAGAGAG
 GAATGGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTA
 GGAGGAAATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGG
 GAAGTTTGGGATATACCCCAAAGTCCTCTACCCCTCACTTTTATGGCCCTTTCCCTAGATA
 TACTGCGGGATCTCTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATA
 TATTTGGAAATTAAAGTTTCTGACTTT

FIGURE 242

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812
><subunit 1 of 1, 337 aa, 1 stop
><MW: 37668, pI: 6.27, NX(S/T): 1
MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPA SYPECGNNAQSPIDIQTDSVTFDPDLP
ALQPHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQ
INSEATFAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRH
KDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTL
FSTEEEPSKLLVQNYRALQPLNQRMVFASFQAGSSYTTGEMLSLGVGILVGCLCLLLAVYF
IARKIRKKRLENRKSVVFTSAQATTEA
```

Important features of the protein:**Signal peptide:**

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 243

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTC
GTGGACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGG
ATCAACTCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTC
CGGATCAGGGAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTG
ATACCATTAACACAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGG
AATGACACCTGGTACCCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAAC
TGCACCCACATGTGTTACCAATTTTTGTACACAACCTTGGAGCCCAGGGCACTATCCTAAGC
TCAGAGGAATTGCCACAAATCTTCACGAGCCTCATCATCCATTCCTTGTTCCCGGGAGGCAT
CCTGCCACCAGTCAGGCAGGGGCTAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAG
GAGCAGGTGTAAATCCTGCCACCCAGGGAACCCAGCAGGCCGCCTCCCAACTCCCAGTGGC
ACAGATGACGACTTTGCAGTGACCACCCCTGCAGGCATCCAAAGGAGCACACATGCCATCGA
GGAAGCCACCACAGAATCAGCAAATGGAATTCAGTAAGCTGTTTCAAATTTTTTCAACTAAG
CTGCCTCGAATTTGGTGATACATGTGAATCTTTATCATTGATTATATTATGGAATAGATTGA
GACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATTTACCTGAAAATATTCTTGAAATT
TCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAACAATAATTCAATGGATAAAT
CTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATATTAAACATATTTGGAAA
ACTGGAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 244

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTLISSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

Sequence = G E E L L L

FIGURE 245

GAGAGAGAGGCGCGGGTGAAAGGCGCATTTGATGCAGCCTGCGGCGGCCCTCGGAGCGCGGC
 GAGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCC
 GCAGCCGGGAGGCCATGCGACCCCAAGGGCCCCGCCGCCTCCCCGCAGCGGCTCCGCGGCCTCC
 TGCTGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAG
 CAAAAGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGG
 GCCAGCAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTG
 GGATCCCAGGTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAG
 GAGTCTGGACACCCAACTACAAGCAGTGTTCATGGAGTTCATTGAATTATGGCATAGATCT
 TGGGAAAATTGCGGAGTGTACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGT
 TCAGTGGCTCACTTCGGCTAAAATGCAGAAATGCATGCTGTCTAGCGTTGGTATTTTACATTC
 AATGGAGCTGAATGTTTCAGGACCTCTTCCCATTTGAAGCTATAATTTATTTGGACCAAGGAAG
 CCCTGAAATGAATTCACAATTAATATTTTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAG
 GAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGGTGGCACTTGTTTCAGATTACCCAAAA
 GGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAATA
AATGCTTTAATTTTCATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCACTTAAAT
 GACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGA
 CCAAAGTGTGATTTTCACTGTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAAGT
 GGTTTCAATATTTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTA
 TAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTAAAAAATA
 TAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTTATATCTGTAAAT
 AAAAAATTATTTCCAACA

FIGURE 246

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPA
PGRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSP
EMNSTINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.